UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

UTILITY AND EUDAIMONIA:
Ethics and Well-Being in the Aristotelian and Utilitarian Moral Traditions

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
Degree of
DOCTOR OF PHILOSOPHY

By
NOAH JONES
Norman, Oklahoma
Fall 2025

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/341753

UTILITY AND EUDAIMONIA

A DISSERTATION APPROVED FOR THE
SHYAM DEV PATWARDHAN DEPARTMENT OF PHILOSOPHY

BY THE COMMITTEE CONSISTING OF:

Dr. Stephen Ellis, Chair

Dr. Russell Jones

Dr. Zev Trachtenberg

Dr. Luis Cortest

©Copyright by NOAH JONES 2025
All rights reserved.

Table of Contents

Abstract vii

Acknowledgements viii

Preface 1

The Relationship of My Work to Ordinary Life 1

The Relationship of My Work to Relevant Scholarship 4

1. Aristotelianism and Utilitarianism as Rival Moral Traditions 7

I. Traditions of Moral Inquiry 7

II. The Aristotelian Tradition 11

Aristotle 12

Aquinas 14

Modernity Against Christian Aristotelianism 16

MacIntyre 17

The Aristotelian Tradition's Commitments 19

III. The Utilitarian Tradition: History 24

Utilitarian Origins: Quarrelling About Genealogies 24

Bentham 27

Mill 32

Sidgwick 36

Post-Classical Utilitarianism 41

IV. The Utilitarian Tradition: Moral Inquiry as Such 44

Introduction 44

Rational, Empirical Ethics 44

Utilitarianism as a Tradition 48

The Challenges of Value Plurality and Criticism 53

V: Conclusion: Utilitotelianism? Aristotilitarianism? 54

2. Irreducible but Commensurable: Moderating Aristotelians and Utilitarians on Value Conflicts 56

Introduction 56

I: Extreme Incommensurabilism (*After Virtue*) 57

Three Interpretations of Tragic Conflict 57

II: Extreme Commensurabilism (Hedonistic Utilitarianism) 62

III: Moderate Commensurabilism (Ellis) 64

Practical Reason Despite Incommensurability? 67

The Main Problem with the Main Argument 68

The Nature of Commensuration 70

IV: Moderate Commensurabilism (*Whose Justice? Which Rationality?*) 72

Apparent Incommensurability as Distinctly Modern 75

V: Analysis 76

A Concluding Proposal 79

3. Aristotelianism, Utilitarianism, and the Gap Between Apparent and Actual Value 84

Introduction 84

I: Aristotelian Perfectionism: A Complex Good 85

II: Utilitarianism Against Aristotelianism 88

IIa: Hedonism: A Simple Good 89

Problem 1: Valuable Things Besides Pleasure 91

Problem 2: Non-Valuable Pleasures 93

Problem 3: Empirical Inaccessibility 94

Hedonism: Upshots 95

IIb: Preferentism: Goodness Formalized 97

Felt Desires and True Ends 98

From Desire to Preference 99
Synchronic Preferentism 102
Diachronic Preferentism 102
Idealized Preferentism 105
Two Sorts of Utilitarian 111
Preferentism: Upshots 116
III: Conclusion: The Tension of The Gap 116
Afterword: What I Have and Haven't Done 121
More Histories 121
More Conflict Points 123
Bibliography 127

Abstract

My dissertation launches a broader project aimed at advancing the philosophical dialectic between Aristotelianism and utilitarianism, arguing with respect to three specific issues that their respective central commitments are not essentially opposed, and their respective central challenges can be better addressed by attention to their apparent rival's insights.

My Preface briefly contextualizes my project within the concerns of normal people and within recent scholarship. In Chapter One, I explain the MacIntyrean conception of "traditions of moral inquiry" that underlies the project, propose a working outline of the central commitments and challenges of Aristotelianism and utilitarianism, then moderate the two traditions' conceptions of moral inquiry as such. Chapters Two and Three address two of the most important conflict points between the traditions—value plurality and the criticism of human dispositions. In both cases, I argue that the two traditions can share a middle ground between untenable extremes. In my Afterword, I describe what I take myself to have accomplished and outline what most urgently needs to be accomplished next in the Aristotelian-utilitarian dialectic.

Acknowledgements

To Steve, I owe the origin, maintenance, and clarification of my project over its several years. It was his observation, uttered casually in the philosophy lounge early in my tenure at OU, that sophisticated forms of utilitarianism often start looking suspiciously like eudaimonism or Aristotelianism. That observation, along with our study of MacIntyre, launched my inquiry and has not lost its intrigue. Steve supported and encouraged me through my many moments of doubt as to whether I had anything interesting, important, or plausible to say. To the extent that I did have such things to say, they largely emerged in conversation with him.

To my committee—Rusty Jones, Zev Trachtenberg, and Luis Cortest, I owe guidance on and improvements to the project of all sorts, ranging from structural clarifications, research directions, fact-checking and concept-checking. Perhaps most importantly, each encouraged me to persist at crucial points.

To my students from the past five years of classes, I owe much of the development of my thought about ethics broadly, and even about utilitarianism and Aristotelianism in particular. They often ask excellent questions that provoke me to deeper thought, and they keep me grounded in ordinary life and conversation.

To the employees of the OU Library, I owe (of course) my access to essential research materials, but especially to the dozens of obscure works by and about Alasdair MacIntyre procured by the Inter-Library Loan department.

To many friends and family members, I owe so much encouragement, as well as patience in listening to dozens of ill-structured and perhaps ill-timed ramblings on the subject.

To Natalie, I owe so very much encouragement, and again, so very much patience in listening to *hundreds* of ill-structured and perhaps ill-timed ramblings. You keep my attention directed toward the tasks most worth doing, and you're part of what makes those tasks worth doing. If I've achieved anything in this project, it's been empowered by your admiration and affection, given independently of my achievement.

Preface

The Relationship of My Work to Ordinary Life

Philosophy dissertations are generally supposed to achieve something within very specific scholarly conversations about abstract ideas. But the concerns of this project are not far away from the concerns of ordinary people, and I find it important, at the outset, to briefly spell out why.

In the following pages, I say a lot about happiness, morality, and rationality. Everyone wants to live a good life, and we suspect that living well has something to do with happiness, rationality, and morality. The three ideas seem related in some important way but also not the same. The trouble with these targets is that all three are quite hard to hit independently, and they're especially hard to hit all at once. There are all sorts of reasons why.

First off, I want to be happy, moral, and rational, but I'm not exactly sure what happiness, morality, and rationality are! There are many different views about what they are. Which should I aim at? What if I can't do what's moral or rational even once I know what it is? Another problem: happiness seems to involve getting all sorts of good things and avoiding all sorts of bad things, but how can I make decisions that account for these different pushes and pulls? And then, will getting what I want, or what feels good, even make me happy?

And ultimately, aren't all these questions so hard that I may as well give up?

Unfortunately, if I give up on answering the questions, I'm really accepting one specific answer—accepting it without thinking hard. To say that attending to ethical questions isn't *worth it* is itself an ethical claim. It implies that I ought to spend my time and attention on something other than inquiring. Suppose I just try to "go with the flow"—act on the impulse of the moment, or follow the social norms of the people around me. In principle, it could turn out that one of these approaches works out better for me than any alternative. But it could turn out otherwise. And if I give up on questioning, my head is in the sand. The alternatives never have a chance.

And in any case, hard questions still show up for these approaches—what do I do when social norms or my own instincts conflict with each other? Because I have to make choices one way or another, moral questions are just as unavoidable as they are challenging. Whether I like it or not, I have to inquire—at least a little. How should I go about it? Given that inquiry is very hard, I'll have to approach it like I approach other hard tasks.

First, I may need help from other people to come to better conclusions about moral issues. It is just as fruitless to try and *think about ethics* by myself as it is to try and *be ethical* by myself. In fact, it is literally impossible to think about ethics entirely on my own, because "my own" thoughts have been shaped, and continue to be shaped, by other people. So *we* need to engage in "ethical inquiry" together—experiencing, interpreting, discussing, and deliberating, and disagreeing together. This might turn out to be very hard. But what other choice do we have? Again, if we give up, we're pretending to be something other than we are—*social* beings. Naturally, the people closest to us will be our most practically important companions, but as our conversation develops it may extend across the world, e.g. through reading and writing. It may also turn out that some of the people who can contribute to our inquiry are dead.

Which leads to a second insight: it may take us a very long time to figure things out. So, as an individual, I'll have to think of ethical inquiry as a long-term project, and I'll need to become a "lifelong learner" pursuing it. Then, as a community of inquirers, we'll have to think of our inquiry as an even longer-term project, extended across generations. We'll study how similar questions were asked and answered by the inquirers before us, and maybe make some progress for the inquirers after us.

This means, finally, that our inquiry will always be open-ended—I might later figure things out that I presently can't, and *we* might later figure things out that *we* presently can't. In fact, we may be better off aiming not to "figure things out" and "make right choices" but to at least grasp things better than before, and make choices better than before. Without giving up on *conviction*, we'll have to give up on *certainty*. The Avett Brothers were right: "Turns out we don't get to know everything."

So my initial desire to live a good life and my initial question about how to do it invest me in some tentative commitments before I've even begun inquiring in earnest. Practical philosophy—thinking hard about what we should do—has to be social, historical, and open-ended. I'm a member of a long-term community (maybe more than one community) of inquiry. We can call that kind of community a "tradition." In my practical life, I answer questions with resources provided by my tradition(s).

This dissertation engages with two specific traditions of moral inquiry. I call them "Aristotelianism" and "utilitarianism," but the names aren't as important as their historical origins and conceptual commitments. In Chapter 1, I discuss these histories and concepts,

highlighting their different approaches to moral questions in general. Besides setting up the next chapters, this comparison draws out how each tradition has had to learn to understand moral inquiry in the way I suggested above—historical, social, and open-ended. Then in Chapters 2 and 3, I discuss some more specific moral disagreements between the two traditions. First, how can I make rational choices when faced with radically different kinds of good and bad things? Members of the utilitarian tradition insist that there must be somehow rational procedure for resolving conflicts between different values. Sometimes, they imply that the procedure needs to be highly precise, like a mathematical algorithm or scientific law—think of the economic cost-benefit analysis, which is supposed to work this way. Members of the Aristotelian tradition, by contrast, often find the differences between values so radical that they think the values cannot even be compared—ordinary people sometimes talk like this about "apples and oranges." I try to show that we really *can* compare values in a way that is rational, not arbitrary; we can reason to judgments about the place that various goods should have in our lives. But I also claim that our judgments won't be certain or scientific; we'll have to continually update and sometimes overturn them as we experience the world and inquire about it together.

The second specific question I discuss is whether there is a difference between what we *want* or *like*, and what is good for us. Again, utilitarians and Aristotelians generally disagree about this. Utilitarians tend to think there is no difference, or at least the difference is minor and easy to overcome. My well-being is, fundamentally, the maximization of my pleasure, or the satisfaction of my desires. By contrast, Aristotelians tend to think there is a big difference between what is good for us and our tastes or desires. Simply put, we often like the wrong things, and we may need to *learn* to like the right things. Aristotelians think utilitarians sound naïve—surely what's good is not the same as what feels good or seems good. Utilitarians think Aristotelians sound elitist—who are you to judge what's good for others? I try to show that although there's ultimately no way to avoid criticizing people's tastes and desires, it may be possible to avoid the worst dangers of elitism.

Both these processes—criticizing my desires and tastes, and fitting diverse values together in my life—require the project of moral inquiry I just sketched. They will have to be extended through time, reflected off others, and always in progress. I think these processes capture much of what we're doing when we argue about "ethics" or "morality." A full theory of

ethics or morality would be much *more* than what I've started working out here, but I don't think it could be any *less*.

The Relationship of My Work to Relevant Scholarship

Most broadly, I study *ethics through history*, the process through which we try to understand ourselves, our values, and our common life. More specifically, I aim to help traditions of ethical inquiry understand each other and perhaps progress toward truth. More specifically still, I want the utilitarian and Aristotelian traditions to more fully reckon with each other than they have so far. This dissertation is one component of that specific project: a careful rapprochement between Aristotelianism and utilitarianism with respect to the nature of moral inquiry, the diversity of values, and the evaluation of tastes and desires.

By Aristotelianism, I mean the specific tradition of thinking about the good life and the virtues that self-consciously identifies with Aristotle's practical philosophy, especially as described and represented by Alasdair MacIntyre. MacIntyre himself did not consider himself a defender of a rival to utilitarianism, deontology, or contractarianism called "virtue ethics." While Aristotelianism certainly does foreground characterological questions, it situates virtue within the equally crucial concepts of teleology, community, and eudaimonia. For this reason, utilitarian or consequentialist attempts to foreground virtue concepts are relevant but not identical to the sort of rapprochement I aim at.[1] In Chapter 1, I briefly outline the Aristotelian tradition's history and central tenets.

By utilitarianism, I mean the specific tradition of moral thinking, with its origin in early modern Britain, that makes well-being, or perhaps happiness, the ultimate standard for all moral evaluations. On one hand, my utilitarianism is *narrower* than mere "consequentialism," which may evaluate outcomes with respect to goods other than well-being. On the other hand, my utilitarianism is *broader* than "classical utilitarianism," which is often taken to include a hedonistic value theory and perhaps a particular psychology of motivation. One of the primary tasks of Chapter 1 is to more thoroughly conceptualize utilitarianism as a tradition; to the extent that I succeed, it will be clearer why I have chosen my particular way of identifying the historical/conceptual "through-line."

---

[1] Driver, *Uneasy Virtue*; Crisp, "Utilitarianism and the Life of Virtue". See Discussion in Railton, "Character and Utilitarianism."

There are two sorts of Aristotelians. Some have somewhat causally suggested that utilitarianism or consequentialism might be acceptable if accompanied by a thick theory of value or human nature, and of course, the virtues. For example, Martha Nussbaum once said in passing, "I have no objection to the cost-benefit analysis if you get all the costs and all the benefits lined up."[2] Comments like these both delight and frustrate me; delight, because I think they identify something interesting and important; frustrate, because they rarely follow through regarding what this implies for the relationship between Aristotelianism and utilitarianism. Why continue treating them as rivals—and importantly, *teaching* them as rivals, given the elision of traditional distinctions? Shouldn't we at least require more clarity about, in Plantinga's phrase, "where the conflict really lies"?

On the other hand, there are Aristotelians like Alasdair MacIntyre, Stanley Hauerwas, and of course Elizabeth Anscombe, who somewhat *less* causally suggest that their tradition is fundamentally incompatible with utilitarianism or consequentialism.[3] The Aristotelian allegations against utilitarians are many and serious: defective value theory, denial of moral absolutes, instrumentalization of virtue, ideological use of "moral fictions," and guilt by association with liberalism.

After Virtue's critical treatments of key thinkers since the Enlightenment generated a great deal of heated response. Disciples of each criticized figure emerged with new defenses and interpretations to rebut MacIntyre's "hasty and on occasion somewhat hackneyed" critiques.[4] Many of *After Virtue*'s critics appreciated aspects of MacIntyre's project, including the centralization of the concepts of virtue, teleology, community, and narrative in practical philosophy. They found MacIntyre's revitalized Aristotelianism admirable but not so incompatible with the views of *X* as MacIntyre seemed to think. Such cases have been made, e.g., with respect to Kant,[5] Hume,[6] and Nietzsche.[7] Even emotivism received a plausible revision and defense with sympathy for MacIntyre's vision.[8] Kierkegaard received a whole volume.[9]

---

[2] Warburton and Edmonds, *Philosophy Bites*.

[3] Anscombe, "Modern Moral Philosophy"; Hauerwas, "Can Aristotle Be a Liberal?"

[4] Scheffler, "Review: After Virtue," 445. MacIntyre himself admitted *After Virtue*'s "incompleteness, the sketchiness of the argument at various points and a good deal of untidiness." MacIntyre, "Moral Rationality, Tradition, and Aristotle: A Reply to Onora O'Neill, Raimond Gaita, and Stephen R. L. Clark," 448.

[5] O'Neill, "Kant After Virtue."

[6] Baier, "Civilizing Practices"; Baier, "MacIntyre on Hume."

[7] Conway, "After MacIntyre: Excerpts from a Philosophical Bestiary."

[8] Waller, "The Virtues of Contemporary Emotivism."

[9] Davenport and Rudd, *Kierkegaard After MacIntyre*.

Relatively few utilitarians have engaged directly with MacIntyre's work, and even fewer have engaged sympathetically. An essay by Paul Kelly is exceptional (in both senses) on this point, sketching out how a more sophisticated utilitarianism might survive MacIntyre's critique after taking it quite seriously.[10] Kelly's piece is insightful but brief, leaving off just where, in my view, the dialectic becomes interesting. John Perry takes the Aristotelian-utilitarian dialectic a bit farther than Kelly—he admits that "eudaimonism" and utilitarianism fit into the broad overarching category of "consequentialism," but ultimately rejects utilitarianism in favor of eudaimonism.[11] His well-argued grounds for doing so map quite directly onto the three central issues I address in the following chapters.

Despite a relative lack in *direct* utilitarian response to MacIntyre, several developments in utilitarian literature bear on this project. Regarding Chapter 2, it has become less clear that utilitarians must have a reductive account of value. On the historical side, Mill has been convincingly read as a value pluralist; on the contemporary side, various thinkers have broadened utilitarianism beyond value monism, including Amartya Sen, James Griffin, Derek Parfit, Shelly Kagan, Thomas Hurka.[12] Regarding Chapter 3, many utilitarians (suspiciously, many of the ones just cited) are willing to criticize human dispositions, though the particular critiques differ. Given that many thinkers from both sides of the divide consider these issues watersheds, it would seem that most philosophers seem not to have noticed or at least adequately appreciated the extent to which these more sophisticated utilitarian accounts of well-being erode the grounds on which utilitarians and Aristotelians often distinguish themselves from each other.

---

[10] Kelly, "MacIntyre's Critique of Utilitarianism."

[11] Perry, "Where Did Utilitarianism Come From?"

[12] Sen, *Development as Freedom*, 76–81; Griffin, *Well-Being*, chap. 5; Kagan, *Normative Ethics*, chap. 2; Hurka, *Perfectionism*, chaps. 6–8; Parfit, *Reasons and Persons*, app. I.

1. Aristotelianism and Utilitarianism as Rival Moral Traditions

This chapter constructs an account of utilitarianism as a *tradition of ethical inquiry*. In Parts I and II, I introduce the MacIntyrean concept of traditions of moral inquiry and sketch an account of Aristotelianism as such a tradition to exemplify the MacIntyrean method and position Aristotelianism for critical comparison with utilitarianism. In Parts III and IV, provide a working interpretation of utilitarianism's history and core commitments and then offer a critical discussion of those commitments with an eye toward my two subsequent chapters.

I. Traditions of Moral Inquiry

My project relies on a particular way of understanding moral philosophy. Across his significant body of work, the late Alasdair MacIntyre developed a "historicist" method of inquiry, crediting Vico, Hegel, Collingwood, and John Henry Newman as ancestors.[13] MacIntyre also demonstrates the method's continuity with the development of dialectic through Socrates, Plato, Aristotle, and Aquinas.[14] For MacIntyre, it is only by this method that rational assessment of moralities and moral philosophies is possible.[15]

Central to the approach is the concept of a *tradition*, which is a particular kind of historically-extended inquiry, argument, or conversation.[16] This kind of conversation centers around certain fundamental agreements—beliefs, practices, institutions, or texts which originate as historical contingencies and function as given authorities for the participants in the conversation. These fundamental agreements, as well as the processes by which they are transmitted through time, underlie the conversation's continuity.[17]

In this project, I treat Aristotelianism and utilitarianism as rival traditions. In both traditions, authoritative figures advance the "conversation" by developing and defending certain core commitments. In the following sections of this essay, I propose interpretations of these core commitments and the challenges they generate. Briefly, I say that Aristotelians are committed to

---

[13] MacIntyre, *WJWR*, 350. The most relevant works on this point are Vico's *New Science* (3rd ed., 1744), Hegel's *Phenomenology of Spirit*, Collingwood's 1946 *The Idea of History*, and Newman's 1845 *Essay on the Development of Christian Doctrine*. More broadly, *Whose Justice? Which Rationality?* attempts to show how MacIntyre's conception of theoretical rationality maintains continuity all the way back through Thomas Aquinas, Aristotle, Plato, and Socrates.

[14] This is one aspect of the argument of MacIntyre, *WJWR*, chaps. 5–8.

[15] MacIntyre, *WJWR*, 354ff.

[16] MacIntyre, *WJWR*, 12. "Conversation" is my preferred term; MacIntyre often uses "enquiry", "argument", or "debate".

[17] MacIntyre, *WJWR*, 12.

seeing ethics as inquiry about *eudaimonia* (literally "having a good spirit," usually translated "happiness" or "flourishing"), an account of *eudaimonia* as excellent rational activity, and the maintenance of a community directed toward that same excellence. Utilitarians are committed to seeing ethics as inquiry about the impartial achievement of happiness, and a conception of happiness as a universally desirable good.

A tradition's commitments are refined over time through conflicts of two kinds. *Internal* conflict is disagreement among a tradition's participants about how to interpret its shared commitments or apply them to new phenomena. For example, Christians in late antiquity disputed how to interpret their fundamental agreement that Jesus is somehow divine,[18] and Marxists disputed the relation between theory and action in the transition from capitalism to communism.[19] Aristotelians, for their part, diverge in their accounts of *eudaimonia*, both in how they identify the characteristic human activity and in their associated schema of virtues. And the utilitarian commitment to the impartial maximization of happiness may express itself in diverging accounts of happiness and the relationship between morality and practical reasoning.

Unresolved internal conflicts may simply dissolve a tradition—for example, MacIntyre takes Hume to have demonstrated that 18th-century Scottish Presbyterianism was unable to coherently maintain its core commitments and practically sustain its ways of life.[20] Alternatively, internal conflicts may result in schism, generating distinct traditions with apparently irreconcilable commitments. Naturally, this is how the schisms between Eastern and Western Christianity, and between Protestant and Catholic Christianity, are often interpreted. Western philosophy's analytic/continental divide may be interpreted along these lines.[21] For my part, I interpret the utilitarian tradition as one of many grandchildren of late medieval Christian Aristotelianism. I say 'grandchildren' to indicate the significance of two distinguishable transitions in English moral philosophy between, say, Grotius and Bentham. Roughly, the first generation begins with Christian natural law ethics and ends with Hobbes, who deprives natural law ethics of a teleological and social account of human nature. In the second generation, British moralists attempt to escape Hobbes' amoral conclusions by rejecting certain of his assumptions,

[18] Pelikan, *The Christian Tradition*, 1: The Emergence of the Catholic Tradition (100-600), chaps. 4–5; Newman, *The Arians of the Fourth Century*.

[19] Knight, *Aristotelian Philosophy*, 108–10; MacIntyre, *ECM*, 280–81.

[20] This is one aspect of the argument of MacIntyre, *WJWR*, chaps. 12–16, esp. 15.

[21] Mahfuz, "Understanding the Division Between Analytic and Continental Philosophy."

including empiricism and psychological egoism. From the fracturing medieval tradition, utilitarians inherited both a moral egalitarianism and a teleological basis for practical reasoning.

The second kind of conflict is *external*. Several outcomes may result from encounters between members of a tradition and those outside it who reject its fundamental agreements. Perhaps the least interesting of these is a stalemate, in which one tradition's commitments are untranslatable to another, or their standards are incommensurable with another's.[22] Since Aristotelianism and utilitarianism did genuinely become distinct and rival traditions, their conflicts ought to be regarded as external conflicts. But a stalemate generated by conceptual untranslatability or incommensurability seems the least likely outcome of their conflict given the relatively recent shared history.

There is one quite particular way in which one tradition may rationally defeat another. It may be that Tradition A can offer a perspective on Tradition B's internal conflicts that is superior to the perspective available given Tradition B's commitments. Importantly, 'superior' here means superior by Tradition *B*'s standards of rationality, not Tradition *A*'s. By adopting this superior perspective, a participant of Tradition B transcends the limitations imposed by their previous commitments. MacIntyre often characterizes developments in the history of science in this way, as in the transitions from Aristotelian to Newtonian physics, then from Newtonian to Quantum physics.[23]

Importantly, Tradition A *cannot* be vindicated over Tradition B by showing that Tradition A's commitments are justified by some standard of rationality to which all traditions or all rational persons must agree. There are no such vindicating standards distinct from all particular traditions. While rival traditions may share some rational standards, such as the authority of logic, that authority is thin and formal, so each competing tradition can use logic equally well to frame their own justifications against others. I take it that MacIntyre's point here is well-expressed by the philosophical commonplace along the lines that one thinker's *modus ponens* is another's *modus tollens*. For this reason, logic alone cannot provide an independent standard by which to rationally evaluate the traditions. Rather, formal logic serves not to establish particular truths, but to clarify relations among possible truths. MacIntyre maintains that it does *not* follow from his view that every tradition has an equal claim on our allegiance,

[22] MacIntyre, *WJWR*, XIX.
[23] MacIntyre, "Epistemological Crises, Dramatic Narrative and the Philosophy of Science."

which would be some version of relativism. Nor does it follow that it's impossible to make genuine claims to truth within a tradition, which would be some version of "perspectivism."[24]

A third possible outcome of encounters between traditions is integration: "two traditions, hitherto independent and even antagonistic, can come to recognize certain possibilities of fundamental agreement and reconstitute themselves as a single, more complex debate."[25] In the years between *After Virtue* (1981) and *Whose Justice? Which Rationality?* (1988), MacIntyre came to see Aquinas as having achieved this kind of integration between Aristotelianism and Augustinian Christianity. The achievement involves showing how each tradition's conceptual core is both compatible with and fruitful in developing its rival's.[26]

Aristotelianism and utilitarianism have now interacted as rival moral traditions for three or four centuries. What is the best way to understand their present relationship? Which of the above outcomes seems the most apt interpretation of their dialectic? Answers to such questions are essentially contestable, but one interpretation may be justified as more adequate than another through both historical and conceptual analysis.

Notice that members of each tradition will take a relatively narrow view of the historical and conceptual core of their tradition. Naturally, a narrow construal of each tradition will render it conceptually incompatible with its rival. Given their commitment to their tradition, they'll suppose that Aristotelianism has been rationally vindicated over utilitarianism, or *vice versa*, in the MacIntyrean sense described above. We might describe such interpretive stances as "fundamentalist" or "conservative" in some broad sense (a sense with no evaluative connotation).

My account of the history and commitments of the utilitarian and Aristotelian traditions suggests, I think, a more ecumenical judgment. Throughout working on this project, I have often wondered whether an Aquinas-style integration between these traditions is possible. It is far too early to defend the feasibility of such a claim, too early to assert the feasibility of a thorough and systematic integration of Aristotelianism and utilitarianism. Any such assertion would have to be

---

[24] MacIntyre, *WJWR*, 351–54.There has been much fruitful dialogue between MacIntyre and his critics on this point. Perhaps the most thorough treatment is Lutz, *Tradition in the Ethics of Alasdair MacIntyre*. MacIntyre, "Relativism, Power, and Philosophy"; MacIntyre, "Moral Relativism, Truth and Justification."

[25] MacIntyre, *WJWR*, 12.

[26] MacIntyre, *WJWR*, 205.

justified concretely by a work analogous to the *Summa*, and I am humble enough to admit I cannot produce that work. But my arguments do clear some of the ground for a mutually enriching synthesis, and they indicate what the future *Summa* would need to accomplish.

"Fundamentalist" Aristotelians or utilitarians will disown the positions I propose as either inauthentically Aristotelian or inauthentically utilitarian. They will think I've failed to respect Sidgwick's advice to "avoid all hasty and premature reconciliations, and to exhibit fairly the divergence of the different methods without extenuation or exaggeration."[27] They will suppose that if my arguments succeed in any way, it will be as evidence that their tradition is rationally vindicated over its rival, and their opponents ought to simply "convert" in the face of their epistemic crisis.

So, these dialectic observations suggest standards for assessing my work in the following pages. A critique of my theses would have to be justified by (1) a better interpretation of the history of these traditions, in which it is shown that their key texts, figures, moments, and activities imply stricter sets of commitments than I have suggested, or by (2) an argument that the commitments I outlined are not compatible in the ways and to the extents that I've suggested. I should emphasize that this dissertation addresses only three particular points of intersection. In my conclusion, I mention a few others that need addressing next, and it may turn out that a fundamental incompatibility between Aristotelianism and utilitarianism appears in the course of those investigations. I remain open to that possibility.

II. The Aristotelian Tradition

The Aristotelian tradition has already been thoroughly theorized by MacIntyre and many others.[28] In what follows, I briefly trace the Aristotelian tradition's central episodes, from its beginnings in the conflicts of 5th century Athens, medieval Christianization, crisis at the dawn of modernity, and revitalization in the late twentieth century. Naturally, my historical sketch is highly selective, tailored to my aims of (1) exemplifying the development of an ethical tradition, (2) clarifying the features of the Aristotelian moral tradition that were reworked and replaced by Enlightenment utilitarians, and (3) positioning the Aristotelian tradition for constructive and

---

[27] Henry Sidgwick, *The Methods of Ethics*, I.VI.3.

[28] I draw heavily on MacIntyre as both a historical interpreter and contemporary representative of Aristotelianism, but not exclusively or uncritically. Other key interpreters shaping my account include Lear, *The Desire to Understand*; Knight, *Aristotelian Philosophy*; Irwin, *The Development of Ethics: A Historical and Critical Study*; Sidgwick, *Outlines of the History of Ethics*; Nussbaum, "Human Functioning and Social Justice"; Nussbaum, "Mill Between Aristotle and Bentham"; Rorty, *Essays on Aristotle's Ethics*; Shields, *Aristotle*.

critical engagement with the utilitarian tradition. I suggest that in light of its history, Aristotelianism's core commitments include the central claim that practical reason has a final end, named by the word '*eudaimonia*' and characterized by finality, completeness, and sufficiency. Further, Aristotelians accept several substantive claims about the nature of *eudaimonia* as excellent rational activity, and a practical commitment to a community of rational agents.

Aristotle

The Nature of Ethical Inquiry

For Aristotle, ethics is a practical inquiry, the end of which is good action. Thus, agents for whom the conclusions of reason are not motivating cannot engage fruitfully in the inquiry. This is why Aristotle suggests that formal study of ethics should not be undertaken until after experience and maturity have made one capable of following the conclusions of reason. Ethical theory as such is not intended to, and indeed cannot, persuade a thoroughly vicious person to accept its conclusions about the good nor motivate them to act in accordance with those conclusions.

Moreover, the rational justification of those conclusions, even for the proper students of ethics, is not once-and-for-all. As MacIntyre interprets the procedure,

> When Aristotle first poses his question about the nature of the final human end in the first book of the *Nicomachean Ethics*, he considers three views that are rivals to his own, shows in what ways the arguments of each fail, and then argues from significant aspects of those failures toward the conclusion that he takes to be justified.[29]

So inquiry proceeds by (1) stating the strongest possible objections to a given position, then (2) rebutting those objections to whatever extent possible, then (3) making arguments for the position which "show that they satisfy whatever constraints have been imposed by the negative conclusions" from stage (2).[30] Since new objections may appear, it is "a recurrent task for Aristotelians in each historical period to confront new as well as old objections to their central affirmations."[31]

The subject of ethical inquiry, for Aristotle, is the human good as such, or the final end of practical reason. For Aristotle, the claim that *eudaimonia* is the final end of practical reason is

---

[29] MacIntyre, *ECM*, 87–88; MacIntyre likely has in mind esp. Aristotle, *NE*, bk. I.5. See also VII.1.5, 1145b3-7, *Topics* I.1, 10

[30] MacIntyre, *ECM*, 88.

[31] MacIntyre, *ECM*, 88.

something like an analytic claim, a gesture as what '*eudaimonia*' means in ordinary usage.[32] Before any substantive account of *eudaimonia* is offered, the claim bears no argumentative weight. But the very idea that there *is* a final end of practical reason, a single good for the sake of which all else is done, is philosophically substantive. It requires the existence of at least one non-instrumental good and denies that our various activities and inquiries are futile.[33]

Aristotle offers several formal conditions, or standards by which to evaluate accounts of *eudaimonia*. He suggests that *eudaimonia* must be intrinsically good, complete, self-sufficient, and stable.[34] Aristotle appeals to these standards to argue that *eudaimonia* must be a kind of activity, rather than a state (like pleasure) or a disposition (like virtue). That activity will be the human *ergon*, the activity distinctly characteristic of human nature.

The Aristotelian Account: A life of excellent rational activity in community

Aristotle identifies the human *ergon* as the activity of reason. The exercise of reason in political activity is necessary and happy (*eudaimon*) in its own right, but the activity of contemplation most fully exemplifies final goodness, completeness, self-sufficiency, and stability, and thus is the human good in the highest degree.[35]

Virtue is the disposition to act according to reason, and the virtues are dispositions in which the various components of human nature behave according to reason in their distinct domains.[36] Because Aristotle considers virtue a capacity, rather than an activity, he does not take virtue to *be eudaimonia*.[37] Further, the role of luck in flourishing shows that virtue is not even *sufficient for eudaimonia*.[38] But virtue is necessary for and partially constitutive of *eudaimonia* insofar as it simply *is* reason exercised excellently in various domains. In the moral virtues, reason as *phronesis* (practical wisdom) governs and integrates the nonrational parts of the soul, issuing in right practical judgments and actions.[39] Through the intellectual virtues, reason as *theoria* (contemplation) comprehends "what is fine and divine."[40]

---

[32] Aristotle, *NE*, I.4.2, 1095a16-22.
[33] Aristotle, *NE*, bk. I.1, 1094a20.
[34] Aristotle, *NE*, I.7, 1097a15-b21.
[35] This claim of course commits me to a certain view of the "Book X Problem" in interpreting Aristotle's ethics. See Lear, *The Desire to Understand*, sec. 6.8. 309-320; Shields, *Aristotle*, sec. 8.7, 400–406; Kraut, "Aristotle's Ethics," Discussion in Sec. 10, see also bibliography in C.3.
[36] Aristotle, *NE*, I.13.11-18, 1102a34-1103a4.
[37] Aristotle, *NE*, I.8.8-9, 1098b30-1099a7.
[38] Aristotle, *NE*, I.8.15-17, 1099a31-1099b8.
[39] Lear, *The Desire to Understand*, 167ff.
[40] Aristotle, *NE*, I.7.1, 1177a14-16.

Pleasure and Desire as Fallible Signals about Goods

For Aristotle, pleasure and the satisfaction of desires as such are not always good, because what a vicious or akratic (weak-willed) person desires and finds pleasant will not be "naturally pleasant" or desirable.[41] Only the virtuous person's tastes can be taken as a guide to which activities are independently good, and thus also pleasant for and desired by the virtuous. Moreover, pleasures are irreducible to each other, being attached to particular activities. As Irwin explains,

> Aristotle's way of distinguishing pleasures shows that he does not think pleasure is some introspectively uniform state, or that its different sources ('pleasant things') are merely different and interchangeable instrumental means to the same end. The pleasures of dice playing, sunbathing, and music are different in kind, not merely in origin.[42]

Eudaimonia as Requiring an Excellence-Directed Community

For Aristotle, a human life can only go well in a specific kind of community, which is structured so as to realize the human good as such, which includes all subordinate human goods and the activities in which they are pursued.[43] The Athenian *polis* (city-state) exemplified this sort of community partially but only partially. Politics, or political science, is an inquiry undertaken self-consciously within a *polis* about how the best community may be achieved. This is why Aristotle understands ethics and politics as a single project—the realization of the best form of human life.[44]

Aquinas

To treat Aquinas immediately after Aristotle leaves out many figures of great importance for the development of Aristotelian practical philosophy, including medieval Jewish and Islamic Aristotelians, and Aquinas's more immediate Latin predecessors like Albertus Magnus. The only way to justify such a leap, besides considerations of space, is to suggest that Aquinas's work includes and advances the central Aristotelian insights of those predecessors. For my purposes, Aquinas's importance lies centrally in his reconceptualization of (1) the moral community, (2) the virtues, and (3) ethical inquiry as such.

---

[41] Aristotle, *NE*, VIII.14.8, 1154b17-23; see also 1113a31ff, 1174a20ff, 1176b17ff.
[42] Aristotle, *NE*, 390. Irwin cites III.10; X.5, 1147b24-31, 1154a7-b5
[43] Aristotle, *NE*, I.2.4-8, 1094a27-b12; I.7.6-7, 1097b7-16.
[44] Aristotle, *Politics*, Bk. I.

Aquinas conceptualized a form of practical Aristotelianism that did not depend on the Greek *polis per se*.[45] His commitment to universal principles of natural law allowed "the possibility that human beings may develop in more than one way."[46] Each individual is responsible for responding to their access to those principles, regardless of their education.[47] There can be a community devoted to pursuing the human good as such without the inaccessible cultural particularities of the Athenian city-state. Nevertheless, education is necessary to be practically benefitted by moral philosophy of the kind Aquinas offers in the *Summa*.[48]

Aquinas also corrects Aristotle, through Aristotelian conceptual resources, as to the significance and value of human dependence. This has partly to do with Aquinas's Augustinian influence:

> A strong thesis about the inadequacies and flaws of the natural human order emerges… The Augustinian understanding of fallen human nature is used to explain the limitations of Aristotle's arguments, just as the detail of Aristotle often corrects Augustine's generalizations.[49]

But the correction also has partly to do with the human condition of dependence as an animal, unrelated to sin. Aristotle is guilty of "importing into moral philosophy the standpoint of those who have taken themselves to be self-sufficiently superior," leaving him "unable to give due recognition to affliction and to dependence."[50] For Aquinas, human flourishing—successful practical reasoning—depends on other reasoners and on divine grace to an extent that Aristotle did not acknowledge.

Aquinas's *prudentia* (translated "prudence" or "practical wisdom") corresponds to Aristotle's *phronesis*. It is still a kind of master virtue, structuring and appropriately directing each virtue to the choiceworthy in each situation; "[p]ractical wisdom is the principal virtue in an unqualified sense, outranking all the others."[51] But a new virtue, *caritas* ("charity" or "love") becomes the master virtue in a different sense. Love infuses and reorients each other virtue by directing them properly toward God and neighbor.[52] In this respect, too, Aquinas leans into the

---

[45] For a statement of the problem Aquinas faced, see MacIntyre, *AV*, 170–73.
[46] MacIntyre, *WJWR*, 179.
[47] MacIntyre, *WJWR*, 179.
[48] MacIntyre, *WJWR*, 178.
[49] MacIntyre, *WJWR*, 205.
[50] MacIntyre, *DRA*, 7.
[51] Aquinas, *ST* II.I, Q61A2, Trans. Thomas Williams, in Hyman et al., *Philosophy in the Middle Ages: The Christian, Islamic, and Jewish Traditions*.
[52] *ST* II.I, Q62, A2

Aristotelian commitment to human sociality and away from the "self-sufficiency" of the ideal agent.

The challenge of Aquinas's intellectual project was to maintain faithfulness to both the Aristotelian and Christian traditions, which were understood by many to be incompatible. He met the challenge (arguably, of course) by using the conceptual resources of each tradition to both justify the worthy claims, and correct the unworthy claims, of the other. Because he had to synthesize apparently incompatible theses from traditions of quite different origins, Aquinas brings the nature of tradition-bound rationality to the fore. He exemplifies how a conflict of traditions can be overcome *rationally*, and even in the *quaestio* structure of his philosophy, he teaches the essential open-endedness of inquiry in a tradition.[53]

Modernity Against Christian Aristotelianism

What Linda Zagzebski calls "The Big Shift" of early modernity involved scientific, moral, and epistemic revolutions.[54] The most relevant feature of its scientific revolution(s) was the rejection of Aristotelian physics, in particular its teleology.[55] Early modernity generated a mechanistic conception of the physical world in which final causes became entirely reducible to efficient causes and thus irrelevant to scientific explanation.[56] This world-picture left the place of agents' ends ambiguous. On a prominent early modern view, originating in Hobbes and inherited in important respects by Bentham, the only natural end toward which humans act is a hedonic end. The human faculty of reason does not propose or refine our conceptions of ends, it merely supplies the means to given and inalterable ends—survival and enjoyment.

The religious dimension of the Big Shift, associated primarily with the Reformation, challenged the authority of the church to interpret Scripture and the classical tradition, undermining Aristotle's authority in practical philosophy. Whereas Protestants first sought epistemic authority in a kind of Biblical foundationalism, their critical methods soon challenged the authority of Scripture itself; e.g., by Spinoza.[57] Biblicism continued (and continues) in many Christian sects, but among Enlightenment thinkers it tended to make way for either empiricism or rationalism.[58] Both these modes of Enlightenment foundationalism sought to secure consensus

---

[53] MacIntyre, *WJWR*, chap. 10; MacIntyre, *3RV*, chaps. 5–6.
[54] Zagzebski, *The Two Greatest Ideas*, 64.
[55] Mayr, *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*, 48.
[56] Shapin, *The Scientific Revolution*, 28–30; Zagzebski, *The Two Greatest Ideas*, 64–78.
[57] Israel, *Radical Enlightenment: Philosophy and the Making of Modernity 1650–1750*, 447–56.
[58] See, e.g, Wolterstorff, "Locke's Philosophy of Religion," esp. 196-198.

by appealing to principles acceptable to *all rational individuals*—either the self-evident truths of reason and their entailments, or the facts accessible through the senses.

Failure to secure any substantive agreement about universal rational principles finally put in question the authority of reason itself. No guides to action and belief remained except given desires and arbitrary will, so the practical conflicts of the era—religious, moral, or political—were not rationally resolvable.[59] No established institution embodied rational competition between conflicting ideas, and there were progressively fewer shared sources of authority from which to argue. And of course, if practical social conflicts are not resolved rationally, then they will be resolved forcefully.

But from most perspectives, a war of all against all is unacceptable. What thus emerged was an acknowledgement that rivals must attempt to live with each other peacefully despite rivalry. Politically applied, this acknowledgment amounts to classical liberalism—understood as the view that the state must not advance any particular conception of the human good, but only provide whatever bare freedoms are necessary for peace among individuals who each act on their own conceptions of their good.[60]

MacIntyre

As MacIntyre so famously argued in *After Virtue*, even that apparently minimal liberal project failed; purported state neutrality about the good has only generated more irresolvable moral and political conflict. The Kantian and utilitarian moral traditions which emerged to philosophically justify the project turned out to rest on "moral fictions," including *rights* and *utility*. Regarding utility, MacIntyre

> If someone suggests to us, in the spirit of Bentham and Mill, that we should guide our own choices by the prospects of our own future pleasure or happiness, the appropriate retort is to enquire: 'But which pleasure, which happiness ought to guide me?' For there are too many different kinds of enjoyable activity, too many different modes in which happiness is achieved. And pleasure or happiness are not states of mind for the production of which these activities and modes are merely alternative means.[61]

Utility is a "pseudo-concept" insofar as it lack action-guiding content and so can be used to mask the arbitrary ends of the speaker. For MacIntyre, utility must be this sort of fiction given the

---

[59] MacIntyre, *WJWR*, 334–36.
[60] MacIntyre, *WJWR*, 336.
[61] MacIntyre, *AV*, 63–65; MacIntyre, *Utilitarianism and the Cost-Benefit Analysis*.

incommensurability of the various pleasures available to me. Of course, more on this in Chapter 2.

Nietzsche's most admirable insight was to perceive this failure of the Enlightenment and predict its consequences. MacIntyre suggests that we must either accept Nietzsche's conclusions—or else, return to the Aristotelian ethics that the Enlightenment began by rejecting. This return would involve restoring a robust *telos* to practical reason—a *telos* not derived, like utility, from what humans *as they already are* happen to want or enjoy. The *telos* will involve a thick conception of human life as a narrative constituted significantly by the virtues it embodies and the communities it develops from. The rational justification of Aristotelian ethics will be a tradition-bound form of rationality, which escapes the arbitrariness of both parochial fideism and Enlightenment rationalism.

But of course, it is both impossible and undesirable to return to Aristotle naively. It is impossible because the *polis* no longer exists; we need a form of social life that expresses the Aristotelian moral project today. We also must avoid Aristotle's exclusions of some types of humans from the active life of the *polis*, and thus from living the best kind of human life. In *After Virtue*, MacIntyre hoped to construct a teleology that was independent of Aristotle's biology.[62] But in *Dependent Rational Animals*, he acknowledges the importance of our animality for ethical inquiry. So any revival of Aristotelian ethics must be accompanied by a contemporary but compatible biology, psychology, sociology, politics. These inquiries will be Aristotelian in the sense of rejecting a kind of reductionism—the reduction of natural goods to arbitrary aims, the reduction of social relations to manipulative interactions between atomic individuals.[63]

---

[62] MacIntyre, *AV*, 162.

[63] A concerted recent effort towards this project has been made by Rob Koons in his Aquinas Lecture. It may be worth quoting at length from his concluding vision of the human sciences from a quantum-Aristotelian perspective:

> Aristotelian pluralism carves a middle path between atomism and cosmic monism, securing a foundation for the manifest image of human life. It acknowledges the homeliness of the world—a place in which human freedom, agency, and knowledge can exist without threat of nihilism or corrosive skepticism. Modern science seemed to threaten this world with a universal acid of atomistic reductionism, but the implications of the quantum revolution enable us to set the world right again.
>
> The holistic and teleological character of quantum mechanics does not by itself vindicate the reality of teleology and agency at the biological or personal levels. However, it does dramatically change the imaginative landscape of modern science, making the supposition of top-down, formal causation in the realms of chemistry, thermodynamics, biology, and psychology plausible. In fact, the trend of science in the last 50 years has been toward greater differentiation, not unity. Taking these scientific results at their face value means accepting causal agency (understood in Aristotelian terms) at many levels, including the macroscopic level of complete organisms. The idea that there could be a natural and fundamental teleology governing human choices is once again fully credible.

The Aristotelian Tradition's Commitments

Despite Aristotelianism's long and complicated history, certain commitments emerge as central in each of its forms. These commitments face challenges from both inside and outside the Aristotelian tradition. I have chosen to highlight only those commitments and challenges that seem most relevant to the rivalry between utilitarianism and Aristotelianism.

Ethical Inquiry as Such

For Aristotelians, moral inquiry is supposed to achieve clarification of the final end of practical reason for educated practitioners and deliberators:

> Surely knowledge of [the best good] also carries great weight for one's way of life, and if we know it, we are more likely, like Archers who have a target to aim at, to hit the right mark. If so, we should try to grasp in outline at any rate, what the good is, and which is its proper science or capacity.[64]

Aristotle can assume that his pupil would desire and act so as to hit the target once it was identified more clearly; "The lectures of the *Nicomachean Ethics* are meant to be a completion and perfection of the training that has gone before," which is why the lectures could be practically beneficial to their audience.[65]

Aristotle advises these pupils not to expect the study of ethics to produce conclusions with a great deal of practical precision:

> Our discussion will be adequate if we make things perspicuous enough to accord with the subject matter; for we should not seek the same degree of exactness in all sorts of arguments alike, any more than in the products of different crafts. Now fine and just things, which political science examines, differ and vary so much as to seem to rest on convention only, not on nature. But goods also vary in the same way, because they result in harm to many people—for some have been destroyed because of their wealth, others because of their bravery. And so, since this is our subject and these are our premises, we shall be satisfied to indicate the truth roughly and in outline.[66]

As Irwin notes, Aristotle here discourages taking the circumstantial nature of evaluative claims, or diverging evaluations, as evidence that "there is no objective truth about what is just and fine," supposing evaluations to depend on convention rather than nature.[67] Aristotelians do not think of

---

These are, of course, Koons' conclusions, not his premises. Koons, *Is St. Thomas's Aristotelian Philosophy of Nature Obsolete?*, 251–52.

[64] Aristotle, *NE*, I.2.2-3, 1094a23-26.

[65] Jonathan Lear, "Alasdair MacIntyre and Therapeutic Method," in *Learning from MacIntyre* (The Lutterworth Press, 2020), 305–6; See also Alasdair MacIntyre, *Whose Justice? Which Rationality?* (Notre Dame UP, 1988), 124.

[66] Aristotle, *NE*, I.2.3, 1094b12-23.

[67] Aristotle, *NE*, 206.

their inquiry as a method for specifying, in advance, how to proceed through practical difficulties.

MacIntyre's traditioned conception of rational justification is by no means ubiquitous among Aristotelians, but it maintains continuity with Aristotle's sensitivity to the limited persuasive scope, and limited decisiveness of ethical conclusions. Moreover, MacIntyre has argued at length that a traditioned account of rationality emerges from Aristotelianism's history as the best way to make sense of its form of inquiry. The method begins in Socrates as the *elenchus*, which can establish, only and at best, that some substantive thesis cannot be presently refuted by the interlocutors. Plato transcends Socrates through dialectic, positing the forms as *archai* of inquiries, knowledge of which would be *as such*, not merely *perspectival*. Aristotle, in turn, transcends Plato with his own method of dialectic, which includes both bottom-up induction *to* first principles (*archai*) and top-down deduction *from* those first principles.[68] The coherence of the whole scheme is what justifies it to those who use it. But Aristotle seemed to impute to his own conclusions an unjustified sort of finality. Aquinas transcends Aristotle by more clearly allowing the possibility of being transcended by future inquiry:

> Aquinas, in appropriating both traditions, integrating them, and carrying forward what was specific in each, did so by a method which required that his own work should be essentially incomplete… For a while the conclusions of each particular question are integrated into an overall hierarchical demonstrative structure which represents the point so far reached in the movement towards a finally perfected science… yet to each question the answer produced by Aquinas as a conclusion is no more than and, given Aquinas's method, cannot but be no more than, the best answer reached so far.[69]

What is arguably implicit in Aquinas becomes explicit in MacIntyre, with significant inflections from Giambattista Vico, G. W. F. Hegel, and John Henry Newman. This is the way of thinking about traditions sketched in Section 1 above.

Eudaimonia as the Final End of Practical Reason

Aristotelians understand *eudaimonia* as a quality that belongs properly to a human life that is good as a whole and also secondarily to the good activities that constitute that life. To

---

[68] This reading most likely makes Aristotle a sort of *coherentist* rather than a foundationalist with respect to *justification*, but a *correspondence theorist* rather than a coherence theorist with respect to *truth*. Truth has to do with the adequacy of our minds in responding to the world, but the only possible guide to what is in the world is the best available synthesis of our various encounters with it. Of course, such categories can only be applied anachronistically to Aristotle. I believe MacIntyre at least should be read along these lines, but even MacIntyre almost never uses these analytic terms in their technical senses.

[69] MacIntyre, *3RV*, 124.

describe a person's life or one of their activities as good is to identify it as exercising the person's characteristic capacities to accomplish their characteristic function. Further, there is a particular human activity in which it is most appropriate to ascribe *eudaimonia* to the agent; it is of course whichever activity most fully exercises characteristic human capacities and achieves its characteristic function.

The claim that *eudaimonia* has essentially to do with a characteristic activity, or the accomplishment of a function, is essential to Aristotelianism. But I propose that the association of *eudaimonia* with *theoria*, as it appears in Aristotle, is not in the same way essential to the tradition in which he stands at the head. Aristotelians have revised and partially replaced *theoria* with different conceptions of human *ergon*. Aristotelians seem to agree that this highest activity involves the exercise of reason, but they do not fully agree on the domain in which reason is exercised, or perhaps the object toward which it is directed. For Aristotle, perfect *eudaimonia* is the exercise of reason to contemplate the universe, as grasped by a complete understanding of its ordered hierarchy. For Aquinas, it is the exercise of reason to know God; more specifically, to perceive the divine essence through the speculative intellect. For MacIntyre in *After Virtue*, it is reflection on the human good as such in rank-ordering the goods involved in a good life.

Eudaimonia and Community

As to *eudaimonia* and community, the Aristotelian claim is that a good human life is a life structured by specific kinds of relationships to other humans. The flourishing of an individual both constitutes and is constituted by the flourishing of their social environments. These contexts include first the family, then the various associations in which particular goods are pursued, then the broader, distinctly political form of community in which the various particular goods are given a determinate place.[70] The Aristotelian conception of *eudaimonia* thus relates the individual's good to the good of their community. This is true even of Aristotle, who foregrounds the ideal independence of the virtuous man, and it is true to even greater extents for Aquinas and MacIntyre.

But to be an Aristotelian is not just to accept Aristotelian claims about the human moral community; it is also to be practically committed to realizing that community. This is why MacIntyre indicates in the famous conclusion of *After Virtue* and elsewhere that being an

[70] MacIntyre, *DRA*, 129–35.

Aristotelian in modernity is challenging.[71] The challenge is not just that modernity makes Aristotelianism hard to believe in various ways, although MacIntyre certainly believes that. It is because Aristotelianism must take a social form, and it is not obvious what form that can be in modernity.[72]

But regarding the conceptual issues, at least two challenges accompany a communal account of human flourishing. The first is the problem of *exclusion* from flourishing, and this of at least three sorts. For Aristotelians, humans cannot flourish if (1) they have no *polis*, (2) their *polis* is wrongly oriented, (3) they associate with the *polis* but cannot participate in its life.

Exclusivity in various forms is a long-standing issue for Aristotelians, of which MacIntyre has long been acutely aware. MacIntyre's most thorough engagement with the challenge appears in his 1999 *Dependent Rational Animals*. MacIntyre, again drawing on Aquinas, criticizes Aristotle for developing an account of the human good as such which does not include most humans. As to (3), MacIntyre reconceives the *polis* in a way that enables all who belong to it to somehow participate in its life, including even the severely disabled.[73] He argues that Aristotle's disparagement of the rational capacities of women and "natural slaves" are not only unnecessary to his account of the good city but positively undermine it, so we can "rescue Aristotle simultaneously from prejudice and from incoherence."[74] He aims to eliminate Aristotle's dichotomy between practitioners and reasoners, or between craftsmen and politicians.[75] As to (1) and (2), MacIntyre's *polis* gives a place to "the stranger," a human outside the community whose needs nevertheless constitute compelling reasons for helping.[76]

A second broad challenge to Aristotelian social theory is that it may seem to leave too little room for individual freedom, since the human good must be achieved through participation in the *polis*, specifically through social roles in the family and other practices. But for a MacIntyrean Aristotelian, it is only through sociality that individuality becomes possible. Moreover, the *polis* may well have a place for the Bohemian, so to speak. Even such a

---

[71] MacIntyre, *AV*, 262–63, xiii–xiv; MacIntyre, "On Having Survived the Academic Moral Philosophy of the 20th Century."

[72] See also the engagement with MacIntyre in Blackledge and Knight, *Virtue and Politics: Alasdair MacIntyre's Revolutionary Aristotelianism*.

[73] MacIntyre, *DRA*, 135–41.

[74] MacIntyre, *ECM*, 86.

[75] See Knight, *Aristotelian Philosophy*. This is a wonderful book-length treatment of Aristotelian practical philosophy with particular attention to this problem. Throughout, Knight attempts to do justice to the interrelated Aristotelian and Marxist strains in MacIntyre's thought.

[76] MacIntyre, *DRA*, 123–24.

personality will always inevitably be defined out of and in contrast to the roles, norms, and structures that enabled them to become an independent practical reasoner.[77]

Eudaimonia and Virtue

Eudaimonia crucially involves the formation and exercise of the virtues. The virtues enable a possessor to accomplish their function excellently, that is, to successfully perform their role in the communal achievement of goods and the good. This successfully performance is accomplished by the governance of reason to direct the whole individual to their final end, whatever that may be. For Aristotelians, the disposition to act according to reason closes the gap between what seems good and what is good. Virtuous people most desire and enjoy genuine goods.

For Aristotelians, eudaimonia is pleasant despite being irreducible to pleasure. Pleasures attach to activities. Eudaimonia does *feel* good, but it *is* good independently of that fact. Pleasure cannot be the human good as such, nor can the pleasant as such be a guide to action, since only the virtuous person's pleasures are reliably attached to good activities. Likewise, the satisfaction of desires is not good without qualification, since it is only good to satisfy desires as tempered by virtue and directed toward the human good.

A key challenge for Aristotelian views of pleasure and desire is the possibility of *alienation*. A conception of *eudaimonia* is alienated from those to whom it is meant to concern when they do not find *eudaimonia* desirable. Plausibly, Aristotelians have sometimes supposed the good life for humans to be a life that most humans don't in fact desire. MacIntyre critiques modern accounts of well-being on the grounds that they "all require a rejection of the view that there is or could be some standard of right action which could be at odds with what human beings generally want."[78] But at other points, MacIntyre seemed to acknowledge the legitimacy of the charge of alienation, even as applied to Aristotle:

> For the Greeks the connection between the moral life and the pursuit of what men want is always preserved, even if sometimes very tenuously. The desires which the moral life is alleged to satisfy are sometimes a little curious, as for example Aristotle's conception of doing philosophy as the supreme fulfillment of human aspiration. But desire is always kept in the picture.[79]

---

[77] MacIntyre, *AV*, 216–21.
[78] MacIntyre, *WJWR*, 75.
[79] MacIntyre, "Notes from the Moral Wilderness," 60.

Alienation is problematic to the extent we expect that an idea of the good life *for me* to be somehow attractive *to me*.

It is open to Aristotelians to suggest that those to whom their conception of *eudaimonia* is not attractive must not have been properly educated—the fault lies in the desirer, not the suggested object of desire. This response may of course be correct, but it will not persuade or help the objector, and it may insulate the conception from external critique and relevant evidence.

III. The Utilitarian Tradition: History

In this section, I sketch an interpretation of the utilitarian tradition, from its origins in the height of the Enlightenment, increasing sophistication through the classical British utilitarians, and key developments and divergences during the 20th century that shape contemporary discussion. As with the Aristotelian tradition, my history will be highly selective, foregrounding the core commitments and conflicts of utilitarianism as a critical response to Aristotelianism. I emphasize utilitarians' conceptions of moral inquiry and the nature of happiness.

Utilitarian Origins: Quarrelling About Genealogies

The central challenge in narrating any tradition's origins is to accurately draw the distinction between it and its predecessors and rivals. As Perry notes in a rich essay, the birth of utilitarianism is surrounded by a family of problematic narratives. His alternative account makes utilitarianism first a theological project, which Bentham and the other classical utilitarians then radicalized into recognizably modern utilitarianism.[80] This account contrasts with

> the 'Standard Story' told by both utilitarianism's champions (it is the one, true rational ethics that throws off the monkish superstition of all past theories) and by utilitarianism's critics (it is a rejection of ethics and an aberration, soon to die out).[81]

The Standard Story begins with institutionalized Christian natural law ethics, which centrally required the notion of a *summum bonum*. That notion was put in question by the Reformation and subsequent political conflicts, as well as by the "anthropological awareness" that emerged from global exploration. Thus, thinkers like Locke, Kant, and Mill propose alternative ethics for a peaceful and tolerant society without such a shared *bonum.* These theories restrict themselves to a "lowest common denominator" for a pluralistic world: "they will not give you the meaning

[80] Perry, "Introduction," 4.
[81] Perry, "Introduction," 5; Perry, "Where Did Utilitarianism Come From?," 24.

of life, as Aristotle and Aquinas would, but they will tell you not to kill your neighbour merely for being Lutheran."[82]

One polemical spin on the Standard Story, which appears, e.g., in Parfit and Singer, emphasizes utilitarianism's self-conscious repudiation of religious ethics.[83] Utilitarianism casts off ancient superstition as "the firstborn child and heir of the Enlightenment."[84] Religious critics of utilitarianism often accept this estrangement. For both parties, the two views seem to be from different historical epochs—the Christian Era and the Post-Christian Era, for better or worse. Other critics argue that consequentialism's denial of moral absolutes so corrupts it that it is hardly a moral theory at all. According to Williams, Anscombe, and the Catholic new natural law theorists (in different ways), utilitarianism thus typifies the failure of modern moral philosophy.[85]

But as Perry argues, these iterations of the Standard Story draw improper lines between utilitarianism and its rivals, historical or contemporary. Against the contrast between utilitarianism and theology, "[u]tilitarianism was originally a Christian endeavor, and its earliest systematic defenses were works of moral theology."[86] Its founders, including William Paley, "were themselves responding to perceived challenges to natural law," and Paley's *Principles of Moral and Political Philosophy* was a standard university text well before Bentham's *Principles of Morals and Legislation* came to be widely read.[87] In Paley's work, he formulates a principle of utility:

> God wills and wishes the happiness of his creatures. And this conclusion being once established, we are at liberty to go on with the rule built upon it, namely, "that the method of coming at the will of God, concerning any action, by the light of nature, is to inquire into the tendency of that action to promote or diminish the general happiness."[88]
>
> So then actions are to be estimated by their tendency. Whatever is expedient, is right. It is the utility of any moral rule alone, which constitutes the obligation of it.[89]

---

[82] Perry, "Where Did Utilitarianism Come From?," 21.

[83] Parfit, *Reasons and Persons*, 454; Singer, *Rethinking Life and Death*.

[84] Singer in particular credits Darwinian science with eliminating the teleology required by natural law ethics; James Rachels emphasizes the unfortunate (but universal) anthropocentrism of pre-modern ethics. Rachels, "A Short Introduction to Moral Philosophy," 7.

[85] Anscombe, "Modern Moral Philosophy," 17; Perry, "Where Did Utilitarianism Come From?," 23; Williams, "A Critique of Utilitarianism," 150.

[86] Perry, "Where Did Utilitarianism Come From?," 24.

[87] Perry, "Introduction," 2, 4.

[88] Paley, *PMPP*, Bk. 2, Ch. 5.

[89] Paley, *PMPP*, Bk. 2, Ch. 6.

Paley certainly had religious critics, who worried that his view opened the door for radical reformism.[90] Bentham, Mill, and Sidgwick seemingly vindicated that worry, and utilitarians like Peter Singer continue to vindicate it. Even by Mill's time, "theological utilitarianism had been entirely eclipsed" by secular and seemingly anti-religious utilitarianism.[91] But Perry makes the crucial observation that broadly teleological ethics has deep Christian roots in patristic, medieval, and twentieth-century thought.

The Standard Story also seems to mistakenly assume that utilitarianism differs from its rivals and predecessors primarily in being *teleological* rather than *deontological*. Both utilitarians and their critics often accept this fundamental distinction between a pre-modern prioritization of the right and a modern prioritization of the good.[92] Perry suggests this is an historical-interpretive mistake that emerged from a distinct focus on the notion of obligation in the 20th century. Anscombe, seems to suggest that pre-modern theories share "a distinctively Christian-Jewish-Stoic legalism," the context for which was lost in the Big Shift.

But in fact, utilitarianism and its historical rivals all "*share* the assumption that ethics is fundamentally about the good or well-being or flourishing." Perry thinks this assumption is apt of most early modern thinkers besides Kant, including Locke, Hume, Hutcheson, Butler, Paley, Bentham, Mill, and Sidgwick.[93] Accordingly, Perry follows the master argument of *After Virtue* in thinking pre-modern theories share "a Christian-Jewish-Aristotelian eudaimonism," which was fractured but not entirely left behind in modernity. Utilitarians, in particular, held onto the Early modern philosophers, hesitant to suppose an essential human nature and telos, "kept trying to build a puzzle with pieces that would not fit."[94] This generalization holds despite divergent early modern conceptions of the nature and measurement of well-being, the relationship of moral rules to well-being, and moral epistemology and motivation.[95]

---

[90] Le Mahieu, "Foreword."

[91] Perry, "Introduction," 2, 4.

[92] Introductions to ethics (e.g. Sandel's *Justice*) often reinforce this mistake by tabling three theoretical options, such as welfare/freedom/virtue, goals/duties/virtues, action/outcome/relationship, etc. Perry remarks that "[s]uch a division makes good pedagogical sense when it comes to introducing newcomers to complex histories and concepts; however, it can be misleading when it comes to understanding the emergence of utilitarianism." Perry, "Where Did Utilitarianism Come From?," 24.

[93] Sidgwick himself testifies to this truth; see Sidgwick, *Outlines of the History of Ethics*, chap. 5.

[94] MacIntyre also differs from Williams and Anscombe in not predicting utilitarianism's inevitable dissipation. For MacIntyre, it persists because our cultural forms of life embody and reinforce the theory.

[95] Perry, "Where Did Utilitarianism Come From?," 25.

Thus, while Bentham's work constitutes a significant historical break, the nature of that break must be described carefully. Rather than the rejection of religion or the foregrounding of happiness in ethics, "the significant change is that with utilitarianism, ethics becomes an engine for social reform."[96] This feature arguably persists from utilitarianism's origin through thinkers like Singer, who are perhaps motivated primarily by sensitivity to unjust suffering. Moral theory serves the reformist agenda.[97] Paley himself argued against intuitionism that it can only justify existing customs, rather than subject them to critique.[98] Despite expecting largely conservative conclusions from his own method, Paley's mode of argument "destabilized his own position." William Wilberforce—a non-utilitarian Victorian reformer—shared some goals with utilitarian reformers, but worried that "without a substantive conception of their goal, mere 'expediency' would corrupt it."[99] Bentham uses Paley's same argumentative mode in support of radically-reformist utilitarianism, as do Mill, Sidgwick, and later utilitarians like Singer.

The English moralists all defended, against Hobbes, the coincidence between accepted norms and human flourishing. Conservative theological utilitarians distrusted the human capacity to evaluate actions and their consequences as God can. If the status quo conforms to divine guidance, what follows is conservatism. If the status quo does not conform to divine guidance, what follows is non-utilitarian reformism *a la* Wilberforce. Reformist utilitarians saw dire social problems—terrible and preventable obstacles to flourishing—and supposed something must be done with or without divine guidance and with or without the status quo.[100] The standards used to critique their society's norms were intended to be justifiable to all rational beings as such. As such, regardless of their society of origin. As such, they had to be universally-accessible. For British Enlightenment thinkers, the best available source for such claims was the evidence of the senses, especially as formulated in the sciences.

Bentham

Ethical Inquiry as Such

---

[96] Perry, "Where Did Utilitarianism Come From?," 26.
[97] Perry, "Where Did Utilitarianism Come From?," 26.
[98] Paley, *PMPP*, V.
[99] Perry, "Where Did Utilitarianism Come From?," 27.
[100] For a fascinating case that the reformist ambition in utilitarianism was *limited* by the strict empiricist standards to which it held itself, see Burns, "Utilitarianism and Reform: Social Theory and Social Change, 1750–1800," 211,224.

Now, to Bentham. Although Bentham's formulation of the principle of utility has broad scope, he does not intend his ethical inquiries to function as a sermon or moral manual for ordinary deliberators. Rather, he offers his work primarily as practical guidance for social and political leaders, advising them to deliberate as rule utilitarians by making laws that can be expected to maximize happiness:

> It is not to be expected that this process [of calculating expected utility] should be strictly pursued previously to every moral judgment, or to every legislative or judicial operation. It may, however, be always kept in view: and as near as the process actually pursued on these occasions approaches to it, so near will such process approach to the character of an exact one.[101]

Bentham acknowledges that the formal hedonic calculus models the *ideal* practical reasoner, and actual political deliberators will only ever approximate that ideal. But he does have to assume his audience's benevolence—assume that his work will reach powerful individuals who already aim to legislate in the interests of all.

Bentham hopes his inquiry will establish ethics and politics as the moral *sciences*. Only on a sound empirical basis will evaluative claims be free from the apparently arbitrary demands of folk morality and religious dogma. Despite being highly revisionary, such an account of morals may hope to secure agreement from all reasonable men, and greatly improve the conditions of all when put into practice.[102] MacIntyre begrudgingly applauds Bentham for not pretending his theory captures what ordinary people mean by moral terms:

> Utilitarianism sometimes has entangled itself, but perhaps need not entangle itself, in questions about the meaning of such words as 'right' and 'good'. Bentham at least made no pretence that his doctrine was an analysis of what moral agents had hitherto meant in using such words; he proposed it instead as a rational substitute for the confusions and superstitutions [*sic*] of earlier moral theory.[103]

Many utilitarians have indeed followed Bentham in distinguishing moral theory as such from accounts of how people use moral terms. Brandt suggests carefully:

> The locations of ordinary language are not always well adapted to say what on reflection we want to say, or to raise questions which on reflection we want to raise… Reliance on paraphrases of ordinary language certified by our linguistic intuitions is a clumsy and

---

[101] Bentham, *IPML*, IV.6.
[102] Shaver, "Utilitarianism: Bentham and Rashdall."
[103] MacIntyre, *Utilitarianism and the Cost-Benefit Analysis*, 220.

> often misleading method for identifying the concepts we need for the guidance of practical reflection.[104]

Happiness as the Final End of Practical Reason

At the heart of this practical-reasoning-replacement, Bentham offered the standard of evaluation already acknowledged by the theological utilitarians. This standard is the core commitment of the utilitarian tradition: the identification of happiness with the first principle of both practical reason and moral value:

> By the principle of utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency it appears to have to augment or diminish the happiness of the party whose interest is in question: or, what is the same thing in other words to promote or to oppose that happiness.
>
> I say of every action whatsoever, and therefore not only of every action of a private individual, but of every measure of government.[105]

It is happiness, empirically conceived, which justifies claims about what one ought (rationally, morally, or politically) to do.[106]

Bentham seems to have anticipated MacIntyre's critique of emotivist moral culture by exactly two centuries. Bentham's utility principle is intended to give sense to moral notions that would otherwise lack it. In fact, a few passages from his corpus seem to suggest that moral imperatives often simply express the speaker's preferences. Bentham asks the proponent of any rival moral principle to

> examine and satisfy himself whether the principle he thinks he has found is really any separate intelligible principle; or whether it be not a mere principle in words, a kind of phrase, which at bottom expresses neither more nor less than the mere averment of his own unfounded sentiments; that is, what in another person he might be apt to call caprice?[107]

---

[104] Brandt, *A Theory of the Good and the Right*, 9–10.

[105] Bentham, *IPML*, I.II.

[106] The final end of *impartial* practical reason is a maximum of the summed happiness of all beings which experience pleasure and pain. Postema and Rosen have collected some evidence that Bentham is "committed not to aggregate net happiness…but rather to the greatest equal happiness." Postema, "Bentham's Equality-Sensitive Utilitarianism." Notwithstanding, Bentham clearly thinks that sometimes, inequality is the best way to maximize happiness and is thus preferable to equality. It is unclear where Bentham draws the line. Shaver, "Utilitarianism: Bentham and Rashdall."

[107] Bentham, *IPML*, 1.14.3.

Some interpreters go so far as to classify Bentham as an emotivist or proto-emotivist. But Bentham likely just means that *unless utility underlies moral imperatives*, they have no content besides expressing the speaker's preferences.

The tension itself is noteworthy for its resonance with MacIntyre's discussion of emotivism as simultaneously (1) a true description of how evaluative language was used in English in a particular time and place, (2) a false theory about evaluative language as such, since language might yet come to have content by connection with teleological standards. Bentham hopes that facts about the *good* (pleasure) will allow judgments about the *right* to transcend mere emotive expression.[108] Despite having a radically different conception of the human good, MacIntyre agrees. Bentham, of course, would reject MacIntyre's thick conception of human nature as too contestable and insufficiently empirical.

Whence Bentham's hedonism about well-being? He apparently thinks his principle of utility needs no proof or argument. Since agents rely on the principle implicitly in action, the burden of proof is on anyone who rejects it.[109] To those who suggest that some pleasures are bad, Bentham replies that such pleasures only seem bad because they regularly cause pain to others. Any potential exceptions we recognize to the principle will have to be justified, either by reference to some good besides pleasure, or to some decision procedure besides acquisition of goods. But Bentham argues that any such purported goods and procedures are approved/used ultimately because they promote pleasure.

I will argue in Chapter 3 that Bentham's form of argument cannot establish principles like Bentham's as simultaneous *incontestable* and *nontrivial*. Practical reasoning may admit of some tautologies or near-tautologies, like "seek the good and avoid the bad." These are indeed incontestable, and agents do indeed rely on them implicitly whether they know it or not. By contrast, there are contestable claims about the nature of good and bad, which turn such tautologies into actionable principles. Bentham's hedonism is such a contestable claim, generating a contestable principle: "seek the pleasant and avoid the painful." As I explain, it cannot be shown that all agents practically rely on this principle without begging the question.[110]

---

[108] MacIntyre, *SHE*, 232–34.

[109] Bentham, *IPML*, bk. I.XII.

[110] Apparently, Bentham may have become aware of challenge and reduced his psychological claim to a mere generalization, adequate for practical political purposes. See Crimmins, "Jeremy Bentham," sec. 3.1.

Nevertheless, Bentham thinks the arguments of those who claim to approve of something independently from its hedonic value are contrived ways of avoiding appeal to some external standard, "prevailing upon the reader to accept of the author's sentiment or opinion as a reason."[111] Pleasure is meant to provide a standard external to sentiment in the sense that one can justify one's approvals to other the community by reference to it. Moral values are thus matters of fact, to be investigated by observation and legislative experiment. Bentham suggests the "truths that form the basis of political and moral science are not to be discovered but by investigations as severe as mathematical ones, and beyond all comparison more intricate and extensive."[112] And in legislation, "the practice of every nation might be a lesson to every other: and mankind might carry on a mutual interchange of experiences and improvements as easily in this as in every other walk of science."[113]

Bentham identifies various qualities that may belong to pleasures or sources of pleasure, including first intensity, duration, certainty, propinquity, then adding fecundity, purity, and extent.[114] The intensity and purity of pleasures aren't quantifiable, but they are commensurable.[115]

In *IPML*, Bentham accepts Hobbes' psychological egoistic hedonism. Pain and pleasure are our masters; in particular, the force that moves us to act is a reaction to *imagined outcomes*, the anticipation of future pleasure and the fear of future pain. These forces are morally neutral, not discriminating amongst feared or desired outcomes. Even benevolent actions have some egoistic base:

> In every human breast, rare and short-lived ebullitions, the result of some extraordinary strong stimulus or incitement, excepted, self-regarding interest is predominant over social interest: each person's own individual interest, over the interests of all other persons taken together.
>
> In the few instances, if any, in which, throughout the whole tenour or the general tenour of his life, a person sacrifices his own individual interest to that of any other person or persons, such person or persons will be, a person or persons with whom he is connected by some domestic or other private and narrow tie of sympathy.[116]

---

[111] Bentham, *IPML*, II.14.
[112] Bentham, *IPML*, Preface.
[113] Bentham, *IPML*, XVI.4.50.
[114] Bentham, *IPML*, 4.2-4.
[115] Bentham, *IPML*, 4.5.
[116] Bentham, *The Book of Fallacies: From the Unfinished Papers of Jeremy Bentham*, 9.3-4.

Bentham aims to reform social policy in such a way that egoistic motives result in altruistic action, as Bernard Mandeville had suggested they already did.[117]

Perry suggests that "[t]he classical utilitarian trajectory launched by Bentham is but one thread in a rich tapestry and not a particularly attractive thread either." Perry specifically objects to Bentham's conception of morality as mathematical rationality, value monism, and separation of morality from theories of human nature.[118] I discuss the first of these in section 4 of this Chapter, and the latter two in Chapters II and III. But it is worth briefly noting how these three distinctives hold together. Because Bentham's epistemic standards were those of the Enlightenment, he aimed to justify ethical conclusions by universal rational standards. The best such standards seemed to be those of the empirical sciences; hence morality-as-mathematics. And this conception in turn required the collapse of plural values into a single measurable good—if various goods are irreducible, then no single empirical measurement could justify normative judgments among them. Finally, universally-justifiable and mathematical ethics could not rely on any robust theories of human nature, which are of course highly contestable.

Mill

Ethical Inquiry as Such

Mill is both the best-known utilitarian and the utilitarian whose utilitarian loyalties are most often questioned. Given his upbringing and his status as utilitarianism's most popular defender, Mill had as strong a claim to being a member of the nascent tradition as any.[119] But at many points Mill seems to stretch the doctrines of his father and Bentham either near or past their breaking points. A central explanation for Mill's apparent heterodoxy was his attempt to integrate insights from diverse sources[120]—perhaps traditions—in a way uncharacteristic of Bentham, whom Mill thought "failed in deriving light from other minds." Many readers take it that Mill failed at this project:

> In the end *Utilitarianism* and *On Liberty*—which, like all his moral works, hide intense intellectual conflict behind a mask of superficial clarity— are expressions of an

[117] MacIntyre, *SHE*, 162–63; Mandeville, *The Fable of the Bees or Private Vices, Publick Benefits*; DiIulio, *Completely Free*, 59.

[118] Perry, "Where Did Utilitarianism Come From?," 29.

[119] Reeves, *John Stuart Mill: Victorian Firebrand*, chap. 1; DiIulio, *Completely Free*, 6 and passim.

[120] DiIulio, *Completely Free*, xi–xii.

> inauthenticity of outlook, which is worth our attention now partly because of the vast numbers of people who have been tempted to share in it.[121]

I direct readers to John Peter DiIulio's ambitious yet persuasive rebuttal of Mill's critics.[122]

> Mill offers a principle about utility as follows:
>
> The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure.

DiIulio argues that neither Mill nor even Bentham offers their utility principle as a *moral directive*. Rather, their formulations assert utility as the foundation for all moral directives; "[w]hat 'actual duties of morality' are to be deduced from this 'creed' is another question entirely."[123] Mill, like Bentham, often situates ethical theory in the context of political deliberation. Individual choices may often be guided by moral rules that make no explicit reference to utility. The function of utility is both to resolve apparent conflicts between moral rules and to explain *why* the actions enjoined by moral rules are, in fact, right.

Mill is an avowed empiricist, requiring all claims, moral and otherwise, to be justified ultimately by experience. Importantly, Mill's empiricism need not be assumed to include a fact/value distinction of kind that Humean or positivist empiricism does. For Mill, empiricism's importance lies in its contrast with intuitionism, which he saw as a failing defense for stagnant conservatism. DiIulio nicely explains the significance of Mill's empiricism for ethical inquiry:

> The challenge for Mill is to demonstrate empirically (without relying on any intuitive, or priori insight) that certain goods or activities can be deemed *objectively* desirable; that at least *some* claims regarding what we either *should* or *should not* take pleasure in—or what we have the *reason* to value or desire—are *more* than just arbitrary, prejudiced assertions. But, frankly, this is a relatively simple task for Mill. We call various things *desirable*, just as we call some apples *red*, because we *experience* them that way: we take pleasure in the thought or idea of pursuing certain goods or activities *because* they are desirable and thus worthy or deserving of our interest; we feel we have *reason* to value or desire certain objects *because* they are enriching, edifying, ennobling. In truth, we have, from one angle, as much occasion to believe in desirability on empirical grounds as we do in red apples; for whereas nothing you do can help *me* see redness through *your* eyes (and nothing you do can *reveal* redness to a blind person), there is a great deal you can do

---

[121] Scruton, *A Short History of Modern Philosophy*, 225–26.
[122] DiIulio, *Completely Free*.
[123] DiIulio, *Completely Free*, 76.

> to help *my* mind's eye appreciate what *your* mind's eye views as desirable in a particular end.[124]

Mill aspires to a progressive liberal order, justified by form of ethical inquiry that is empirically grounded in this less-than-Humean sense.

So, Mill agrees with Bentham that happiness is the final end of individual and political practical reason.[125] But Mill heavily restructures or even replaces Bentham's hedonistic value theory. In chapter 1, he claims

> the ultimate end, with reference to and for the sake of which all other things are desirable (whether we are considering our own good or that of other people), is an existence exempt as far as possible from pain, and as rich as possible in enjoyments, both in point of quantity and quality…

Notice that Mill identifies the ultimate end of human desire as a certain kind of "existence." The end is not a particular experience at a particular point within a life, as perhaps in a certain conception of the beatific vision. Nor is it a particular *kind* of experience in a life, as contemplation or virtuous activity might be in an "exclusive" interpretation of Aristotle's understanding of *eudaimonia*. Rather, the end is a life that has a certain summative feature, that applies to it as a whole.

In a famous chapter 2 passage which R. M. Hare called "rather unhappy," Mill acknowledges both the variety of goods, and the priority of some goods over others.

> …a beast's pleasures do not satisfy a human being's conceptions of happiness. Human beings have faculties more elevated than the animal appetites, and when once made conscious of them, do not regard anything as happiness which does not include their gratification…. [T]here is no known Epicurean theory of life which does not assign to the pleasures of the intellect; of the feelings and imagination, and of the moral sentiments, a much higher value as pleasures than to those of mere sensation… It is quite compatible with the principle of utility to recognise the fact, that some kinds of pleasure are more desirable and more valuable than others. It would be absurd that while, in estimating all other things, quality is considered as well as quantity, the estimation of pleasures should be supposed to depend on quantity alone.[126]

---

[124]DiIulio, *Completely Free*, 55–56.

[125] Aggregation: Like Bentham, Mill conceives of the impartial good as an aggregate of individual and fungible units of well-being. He attempts to show that such a conception of the good does not preclude strict deontological constraints on individual action for the sake of the good. It is, of course, a long-standing issue whether Mill and other utilitarians can consistently claim robust notions of justice.

[126] Mill, *Utilitarianism*, chaps. 2, 15–16.

While Mill does not think of higher and lower pleasures as reducible to some common currency of value, he does seem to think of them as commensurable or proportionable, in the sense that there is a proper way to distribute time and attention in pursuit of different goods. Mill's inclusive-end conception of happiness appears in his response to the claim that happiness cannot be an appropriate aim, since it is impossible. Mill, suggesting that the claim is "at least an exaggeration," reasons as follows:

> If by happiness be meant a continuity of highly pleasurable excitement, it is evident enough that this is impossible. A state of exalted pleasure lasts only moments, or in some cases, and with some intermissions, hours or days, and is the occasional brilliant flash of enjoyment, not its permanent and steady flame. Of this the philosophers who have taught that happiness is the end of life were as fully aware as those who taunt them. The happiness which they meant was not a life of rapture, but moments of such, in an existence made up of few and transitory pains, many and various pleasures, with a decided predominance of the active over the passive, and having as the foundation of the whole, not to expect more from life than it is capable of bestowing. A life thus composed, to those who have been fortunate enough to obtain it, has always appeared worthy of the name of happiness.

The rational agent will not expect constant exalted pleasure, but will aim to "compose" a life that includes predominance of the best available pleasures, which life will also inevitably include some lower pleasures and pains.

According to Mill, humans have a unique "sense of dignity" that leads us to prefer a life of higher pleasure to one of lower, if we have the wisdom that comes from both kinds of experiences. MacIntyre suggests Mill may be reaching toward MacIntyre's own (Aristotelian) distinction between internal and external goods.[127] Not pleasures as such, but the pleasures of the wise, can be trusted as a guide to genuine value.

Mill's objectivism would seem to open him to charges of elitism that Bentham hoped to avoid. Mill's resources for responding to such a charge include giving liberty and autonomy a constitutive place in his conception of flourishing.[128] This puts Mill squarely between classical and modern ethical thought, retaining both a thick notion of the human good and also an aversion to political paternalism.[129]

---

[127] MacIntyre, *AV*, 199.

[128] Mill, *On Liberty*, chap. 3.

[129] Nussbaum, Annas, and even MacIntyre have all to some degree acknowledged Mill's attempt to accommodate a more classical understanding of happiness or human flourishing than he inherited from Bentham and his father.Nussbaum, "Mill Between Aristotle and Bentham"; MacIntyre, *AV*. See DiIulio, *Completely Free*.

Mill at least enriches, and at most replaces, Bentham's psychological egoism. Mill thinks enlightened self-interest already generally coincides with moral obligation; individual rationality coincides with impartial rationality. But this is not so, as Bentham thought, merely to the extent that a well-legislated society makes it so. Mill gives significance to natural human sympathy, the motive of benevolence by which we take pleasure in others' well-being. The dispositions of sympathy require cultivation, but this is no reason to distrust them. Unlike "artificial" desires and pleasures, those associated with sympathy persist upon reflection.[130]

Sidgwick

Despite the inadequacies of Anscombe's genealogy of consequentialism, she is nevertheless right to give Sidgwick great historical significance. My alternative genealogy of the utilitarian tradition comports with Sidgwick's own interpretation of utilitarianism's origins. For Sidgwick, Hobbes' egoism was the most serious and radical rival to all moral theories, so pre-Bentham utilitarians united with intuitionists (or common-sense morality) against it. The alliance was stable only while the moralists (Shaftsbury, Hutcheson, Hume, Clarke, etc.) assumed the coincidence of common morality with common happiness. But Bentham, "coldly logical" and more open to reformism, "bites the bullet" and admits a gap between the common good and common morality. Where Hobbes suggests I have a natural right to do anything within my powers, Bentham objects by rejecting *all* natural rights.

Sidgwick understood Bentham's innovation in separating utilitarian morality from common morality. But whereas Bentham

> was insufficiently historically aware to know what he was doing… Sidgwick, who *is* historically aware, sees what is going on: Hobbes is the great rival to Benthamite utilitarianism. Not knowing any other way out, Sidgwick opts *intentionally* for the latter, knowing full well the implications for the earlier consensus that had held together common and utilitarian morality.[131]

Perry suggests Sidgwick's dilemma was false and he ought to have known so, being aware of less radical alternatives to Hobbes offered by the pre-Bentham English moralists.[132] I suppose

---

[130] Brandt makes an extended argument to this effect. Brandt, *A Theory of the Good and the Right*, VII.3.
[131] Perry, "Where Did Utilitarianism Come From?," 28n26.
[132] Sidgwick seems to count Hume as a utilitarian, but Hume is certainly not a classical utilitarian.

that despite perhaps being less radical with Bentham, Sidgwick was considerably more reformist than the theological utilitarians.[133]

Sidgwick addressed moral epistemology more thoroughly than either Bentham or Mill. In *Methods of Ethics*, he defends an intuitionist epistemology the standards of which he ultimately fails to meet. Like Aristotle, his methodology involves assembling the best-formulated opinions of the best-educated thinkers, and constructing the best possible synthesis of these facts, which I suggest correspond closely to Aristotle's *endoxa*, the reputable opinions that form the starting points of a dialectical inquiry.

Sidgwick's category of a 'method of ethics' does not map neatly onto contemporary divisions in ethical theory. It is neither simply a moral theory—a fundamental account of what is morally right or good—nor an account of the morally best decision procedure. A method of ethics is a fundamental account of what we have most reason to do. In this respect, Sidgwick's conception of 'ethics' is more nearly pre-modern than modern sense; and he clearly distinguished between the two in *Outlines of the History of Ethics*, noting that ancient philosophers'

> speculations can scarcely be understood by us unless with a certain effort we throw the quasi-jural notions of modern ethics aside, and ask (as they did) not 'What is Duty and what is its ground?' but 'Which of the object that men desire and think good is truly or most desirable, the Good or the Highest Good?'"[134]

Sidgwick recognized that the foregrounding of normative obligation is distinct to modern thought. MacIntyre himself said "[n]o one could write in English on the history of moral philosophy and not feel awed by the example of Henry Sidgwick's *Outlines of the History of Ethics*."[135] In fact, Sidgwick highlighted Aristotle as the model for clarifying and systematizing "our" distinct common-sense morality into a rigorous science of ethics. In light of all this, it remains puzzling that *Methods of Ethics* lumps Aristotelianism and other non-hedonistic forms of eudaimonism in with intuitionism. It is unclear why Sidgwick didn't count Aristotelians in with utilitarians or egoists while acknowledging their disagreements about the nature of happiness.

---

[133] Henry Sidgwick, *The Methods of Ethics*, 2nd ed., 75; Perry, "Where Did Utilitarianism Come From?," 27.

[134] Henry Sidgwick, *The Methods of Ethics*, 2nd ed., 95.

[135] MacIntyre, *SHE*, xxi.

According to Sidgwick, utilitarianism is founded on the axiom of Rational Benevolence: The good of each is equally morally relevant to each agent, though each agent's own good may be more knowable or attainable for them.[136] Utilitarianism transforms this intuition into a method of ethics when it construes each agent's good as happiness:[137]

> By Utilitarianism is here meant the ethical theory, that the conduct which, under any given circumstances, is objectively right, is that which will produce the greatest amount of happiness on the whole; that is, taking into account all whose happiness is affected by the conduct.[138]

Sidgwick was clearer than his utilitarian predecessors in distinguishing between the principle of utility as (a) a deliberation procedure for an individual altruistic agent, and (b) an analysis of impartial practical reason. From this distinction arises Sidgwick's infamous comments in support of an esoteric, self-effacing, or "government house" utilitarianism:

> it may be desirable that Common Sense should repudiate the doctrines which it is expedient to confine to an enlightened few. And thus a Utilitarian may reasonably desire, on Utilitarian principles, that some of his conclusions should be rejected by mankind generally; or even that the vulgar should keep aloof from his system as a whole, in so far as the inevitable indefiniteness and complexity of its calculations render it likely to lead to bad results in their hands.[139]

Several interpretive challenges attach to Sidgwick's account of happiness, which, despite its interpretive difficulties and internal tensions, has been influential on research continuing into the present, including work by Daniel Kahneman, Katarina de Lazari-Radek, and Peter Singer. Sidgwick seems to be a hedonist, claiming that no goods are independent of pleasure. For example, virtue is good, but only instrumentally—virtuous people enjoy acting virtuously or their virtue has beneficial results for others. He also argues that commensuration among specific virtues requires appeal to a hedonic measure, and that commonsense rejections of hedonism are based on misunderstandings of the doctrine or misguided inferences from the paradox of hedonism.

At some points, Sidgwick indicates that pleasure is good because of our attitudes towards it. He equates pleasure with "desirable consciousness,"[140] and he suggests it is not some distinct "tone" of pleasant things than makes them valuable, but rather, as Sumner summarizes, "the fact

---

[136] Crisp, "Utilitarianism in the Late Nineteenth Century," 93.
[137] Sidgwick thinks justice is lexically secondary to utility.
[138] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bk. 4.1.1.
[139] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bk. 4.5.3.
[140] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bk. 3.14.3.

that we like them, enjoy them, value them, find them satisfying, seek them, wish to prolong them, and so on."[141] But Sidgwick *also* probably thinks that desires *per se* are not value guides, even if fully informed, because they may still be irrational or unreasonable.[142]

Sidgwick noticed challenges for the project of "empirical hedonism"—learning from practical experience what produces the most pleasure and pain. Even for one person, it can be hard to rightly judge expected utility. Then interpersonal comparisons considerably complicate the process. Like Mill, Sidgwick acknowledges that common-sense moral rules generally produce useful actions, so they can act as "middle axioms" between the principle of utility and individual deliberation. But these rules of thumb often recommend the wrong actions, and so the challenge of empirical hedonism on the social scale seems unavoidable. Sidgwick thought that human nature is so malleable over space and time that no universal truths of human nature are thick enough to be useful here and now.[143]

The notion of empirical hedonism is a crucial one in that it constitutes a utilitarian admission that it may require a great deal of empirical work to determine how to achieve maximum pleasure. On one hand, this work ought to function like any ordinary science, investigating the *means* to happiness, from what sources happiness can most effectively be derived. But because Sidgwick acknowledges the plasticity of human nature and the complexity of pleasure itself, he commits utilitarians to a dual inquiry, both into *what we enjoy* and *how to effectively get what we enjoy*. Ethics must address both questions of ends and means.

Sidgwick is also ambiguous about whether pleasure is measurable and different pleasures are commensurable. At some points, he seems to deny that pleasure is "a measurable quality of feeling independent of its relation from volition." But again, he apparently needs quantitative commensurability to resolve conflicts between the virtues and other goods.[144]

Besides not necessarily being a good deliberative procedure for the unenlightened, the principle of utility cannot be demonstrated to be superior to the "method" of egoism.[145] Egoistic and altruistic rationalities motives cannot be shown to coincide in every case, despite "the vital

---

[141] Sumner, *Welfare, Happiness, and Ethics*, 86.

[142] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bk. 1.9.3.

[143] Crisp, "Utilitarianism in the Late Nineteenth Century," 95–96.

[144] Sidgwick's account of empirical hedonism may benefit from integration with the "experience-guided calibration" that Ellis proposes for value commensuration. Ellis, "Multiple Objectives"; Ellis, "The Main Argument for Value Incommensurability."

[145] Singer and De Lazari-Radek, *The Point of View of the Universe: Sidgwick and Contemporary Ethics*, chap. 6.

need that our Practical Reason feels of proving or postulating this connexion of Virtue and self-interest." Sidgwick continues in a famously discouraging passage:

> For the negation of the connexion must force us to admit an ultimate and fundamental contradiction in our apparent intuitions of what is Reasonable in conduct; and from this admission it would seem to follow that the apparently intuitive operation of the Practical Reason, manifested in these contradictory judgments, is after all illusory.[146]

Sidgwick thinks of moral reasons as *sui generis*, carrying distinctive motivations which may or may not override others. Sidgwick differs from Bentham and Mill on this point as well as his rejection of empiricism, psychological egoism, and associationism.[147] Sidgwick clearly thinks badly of hedonistic egoism as a method—the combination of welfare hedonism (well-being = hedonism), and rational egoism (rational action results in my well-being). But he finds egoism in irresolvable conflict with utilitarianism, the two being rival accounts "about what it is ultimately reasonable overall to do." For Sidgwick, it is not sufficient if there is merely *some* coincidence between the principles because of natural sympathy and the interest of the whole. Only complete coincidence would resolve the conflict—if maximizing one's own happiness maximizes everyone's happiness, or vice-versa. This is what theistic utilitarians had supposed, but for Sidgwick, even enlightened self-interest was inadequate to resolve the conflict. And Sidgwick's epistemic standards of clarity and consensus could not break the tie between the two methods either.

Sidgwick considers whether to treat the coincidence of motivations as a postulate supported only by our tendency to believe it and its necessity for the systematic coherence of our beliefs.[148] We must either accept it or accept moral skepticism. But such a conclusion would weaken Sidgwick's epistemic standards, one of which is the self-evidence of fundamental principles. Since egoism and altruism are not consistent, they cannot prove self-evident. So Sidgwick admits "an ultimate and fundamental contradiction in our apparent intuitions of what is

---

[146] Henry Sidgwick, *The Methods of Ethics*, 7th ed., "Concluding Chapter" Sec. 5.

[147] Mill perhaps rejected egoism and associationism as well. DiIulio, *Completely Free*, 47–53, 62–66.

[148] Because Sidgwick refuses to evaluate whether divine revelation is to be taken seriously, he cannot hope that God's justice will guarantee the coincidence. He does allow the possibility of theistic ethics as long as the epistemology is intuitionist and the ontology is independent of God's commands, perhaps grounded in God's own end of universal happiness. But he wouldn't justify believing in God merely on practical/ethical/rational grounds; it's not self-evident that there is ultimate justice. "The Sidgwicks [Henry and his wife Eleanor] hoped that their psychical research might ultimately support some core religious claims, particularly concerning the personal survival of physical death, which they deemed a vital support for morality." They were hoping to vindicate the theological utilitarians, who had a clear answer to questions of moral motivation. See Schultz, "Henry Sidgwick," sec. 1.

Reasonable in conduct; and from this admission it would seem to follow that the apparently intuitive operation of the Practical Reason, manifested in these contradictory judgments, is after all illusory."[149] He concludes the first edition of the *Methods* pessimistically with "the Cosmos of Duty is thus really reduced to a Chaos: and the prolonged effort of the human intellect to frame a perfect ideal of rational conduct is seen to have been foredoomed to an inevitable failure." In later editions, he qualifies this pessimism to ensure that nobody concluded that it is reasonable to abandon morality. Frequent coincidence between motives is adequate justification for most choices, but occasions of conflict must be decided "in practice by the weight of the 'non-rational impulses' in play."[150]

So *unlike* Aristotle, Sidgwick does not envision his work as the enduring culmination of all previous inquiry—quite to the contrary, he self-consciously leaves conceptual work for his successors.

Post-Classical Utilitarianism

In the twentieth century, core utilitarian commitments were precisified and applied, and many utilitarians disputed classical utilitarianism's psychological and axiological doctrines. Three lines of development are worthy of brief note here, having to do respectively with utilitarian metaethics, normative theory, and accounts of well-being.

First, during the period in which desire-satisfaction utilitarianism arose, Anglophone philosophy was generally becoming less systematic—the divides widened both between philosophy and other disciplines, and among the branches of philosophy. The most relevant distinction is that between ethical theory and metaethics. Compared to the classical utilitarians, 20th century utilitarians were less committed to particular views about the metaphysics, epistemology, or psychology of morality. But many, including J.J.C. Smart and perhaps Richard Brandt, accepted some form of non-cognitivism—the prescriptivism of R. M. Hare or emotivism of A. J. Ayer.[151]

Second, *indirect* utilitarians make utility the final standard, not for the moral evaluation of individual acts *per se*, but the moral evaluation of *something else*—rules, motives, character

---

[149] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bks. 4, Concluding Chapter, 5, pg. 508.

[150] Crisp, "Utilitarianism in the Late Nineteenth Century," 101.

[151] Smart, "An Outline of a System of Utilitarian Ethics." Brandt disowns non-cognitivism, but I worry that his functionalist account of desire and pleasure arguably involves some loss of normativity similarly to non-cognitivism. See Brandt, *A Theory of the Good and the Right*, I–III.

traits, social policies. Rule-utilitarianism, the most prominent of these, was arguably proposed by George Berkeley and various other utilitarians both pre-classical and classical, including Mill. But in any case, indirect versions of utilitarianism became more prominent after the "Big Three," and the distinction became direct and indirect utilitarianism became sharper when Richard Brandt introduced the act/rule terminology.[152]

The principle that happiness is the final end of practical reason remains as part of the utilitarian inheritance, though some consequentialists prefer 'well-being' to happiness, which suggests a turn away from the more subjectivist understandings of the final end of practical reason. The post-classical utilitarians diverge even more substantially from the classical in another way. Happiness or well-being is not any particular experience, but the satisfaction of desire or preference *as such* either constitutes or is the only possible guide to the nature of happiness.

Desire or preference utilitarians' accounts of happiness range from highly thin and formal to relatively thick and robust. Some desire utilitarians take desires as they are, uncriticized—what we desire is unqualifiedly our good. Other theorists somehow idealize given desires before taking their satisfaction to constitute a person's well-being. Moderate idealizations exclude only inconsistent preference sets and perhaps factually inadequate preferences. More strict idealizations may also exclude preferences with problematic origins, or which are malicious or "bizarre." Less critical accounts aim for incontestability and motivational immediacy; more critical accounts are less neutral but more plausibly adequate to human nature.

In coordination with these philosophical shifts, the nascent human sciences focused attention on motivation, psychology, and practical reason. In particular, the development of rational choice theory should be understood as a component of the utilitarian tradition, the descriptive and empirical counterpart of the normative and conceptual system generally referred to by 'utilitarianism'.

Empirical researchers like Daniel Kahneman have drawn attention to several crucial theses about the nature of practical rationality. Behavioral theorists make it clear that a theory of rational choice can only *explain and predict* an agent's behavior to the extent that it incorporates

[152] Brandt, *Ethical Theory*, 380. As with the transition toward desire-based accounts of well-being, I do not mean to say that act-utilitarianism had or has no defenders. Perhaps the best-published defense of utilitarianism since Sidgwick's was Smart's, which advances act-utilitarianism.

accurate assumptions about that agent's values. The predictive failures of classical expected utility theory resulted from faulty assumptions about what agents care about. Dozens of cases would make this point clear, but consider one from Amartya Sen:

> Suppose the person faces a choice at a dinner table between having the last remaining apple in the fruit basket (y) and having nothing instead (x), forgoing the nice-looking apple. She decides to behave decently and picks nothing (x), rather than the one apple (y). If, instead, the basket had contained two apples, and she had encountered the choice between having nothing (x), having one nice apple (y) and having another nice one (z), she could reasonably enough choose one (y), without violating any rule of good behavior.[153]

If we assume the agent's only relevant goal is apple-maximization, then her behavior looks inconsistent. But as Sen observes, "the presumption of inconsistency may be easily disputed, depending on the context, if we know a bit more about what the person is trying to do."[154]

Empirical choice research can also teach that a theory of rational choice can only *motivate* (provide internal reasons for) an agent to the extent that it incorporates accurate assumptions about that agent's values.[155] This is perhaps intuitive—rational choice axioms have no psychological force until combined with values or ends, apparent goods or bads, that can push or pull agents toward actions. Finally, Kahneman's work in particular highlights that a theory of rational choice can only *obligate* an agent to the extent that it incorporates accurate assumptions about *real* values.[156] This sense of obligation involves the authority of *good reasons*, the kind of external reasons that MacIntyre and Parfit defended against the likes of Bernard Williams.[157] Kahneman raises question about how agents conceive of values—i.e., the ends and final end of practical reason. Specifically, Kahneman's studies amply reveal the irrationality of dis-integrated agency, e.g., when framing effects restrict the scope of values activated in a given choice scenario. Behavior that would otherwise be unintelligible becomes intelligible once the observer distinguishes between various faculties and entities that comprise a human agent—fast and slow thinking, experiencing and remembering self, etc. Psychological dis-integration is crucial to Kahneman's explanatory project. But that project is accompanied by many normative suggestions about how to be *one rational agent*, such as taking up a particular "risk policy" to

---

[153] Sen, "Internal Consistency of Choice," 501.
[154] Sen, "Internal Consistency of Choice," 501.
[155] Kahneman, *Thinking, Fast and Slow*, pt. IV.
[156] The extent to which Kahneman accepts this third claim is not yet entirely clear to me.
[157] Williams, "Internal and External Reasons."

avoid judgment reversals. Though not always explicitly articulated, Kahneman has a normative vision of agency that integrates disparate faculties in pursuit of a coherent rank-ordering of goods.[158]

Of course, despite the academic separation of descriptive and normative projects, utilitarian philosophers have often found psychology relevant to their work, and psychologists have found moral theory relevant to their work. But utilitarianism as a doctrine is no longer understood to commit its defenders to a specific psychology of human motivation, much less Bentham's early egoism.[159]

IV. The Utilitarian Tradition: Moral Inquiry as Such

Introduction

In Section 1, I outlined an Aristotelian conception of moral inquiry (not, of course, the only conception available). In what follows, I develop some theses about utilitarian conceptions of moral inquiry, with an eye to the rivalry with Aristotelianism. I propose that a traditioned conception of moral inquiry is the best way to overcome the challenges faced by utilitarianism's self-conception.[160] To acknowledge the tradition-bound nature of ethical inquiry, with all its limits, would be to acknowledge the failure of some utilitarians' fullest ambitions. But this does not entail that utilitarianism itself failed and fails. Ethics-as-empirical, in its most ambitious form, is not a *core* utilitarian commitment; moreover, a tempered form of ethics-as-empirical may survive in continuity with utilitarianism's history.

Rational, Empirical Ethics

The attempt to found ethics on a rational basis is part of utilitarianism's conceptual core.[161] Utilitarians think ethics is about giving reasons for actions or lifestyles that are distinctly moral.[162] The "moral" point of view is impartial, having equal regard for the interests of each; it is, in Sidgwick's famous phrase, "the point of view (if I may say so) of the Universe."[163] If an action is justified from this point of view, it is justified for all rational beings. For utilitarians in

---

[158] Kahneman, *Thinking, Fast and Slow*, pt. V.

[159] Among others, Kahneman exemplifies the latter claim, and Richard Brandt exemplifies the former.

[160] Both Kelly and Perry have more or less suggested that this would be a fruitful direction for utilitarians. Perry even sketches out some key commitments and historical episodes for the tradition, but he explicitly disavows utilitarianism in favor of "eudaimonism."Kelly, "MacIntyre's Critique of Utilitarianism"; Perry, "Where Did Utilitarianism Come From?"; Perry, "Introduction."

[161] Perry, "Introduction," 4.

[162] DiIulio, *Completely Free*, 107–8; Mill, *Utilitarianism*, chap. 5.

[163] Henry Sidgwick, *The Methods of Ethics*, 7th ed., bk. 3.13.3.

particular, a "rational" basis has generally meant a scientific one. Indeed, making ethics empirical in some sense seems to be a constitutive project of the tradition. Why?

The territory of ethics had always been disputed, but the early modern Big Shift, discussed above, left those disputes especially intractable, not to mention practically urgent. In disputed conceptual territory, all claims must be justified against rivals somehow. Utilitarians rejected two specific forms of justification—appeals to religious authority, and appeals to moral intuition. Both forms of argument are unsatisfactory because the realms of religion and intuition exemplify the same deep diversity as the realm of morality. Some realm independent of these had to ground all appeals within them.

This is why it mattered to utilitarians to make morals susceptible of empirical analysis. Appeal had to be made to something accessible to Catholic and Protestant, rebel and loyalist, perhaps Easterner and Westerner alike. British thinkers characteristically appealed to the world of *experience*, especially as systematized in the sciences. The distinctly *moral* component of the empiricist project was, roughly speaking, to defeat Hobbes' egoism on Hobbes' own epistemology. Thus, ethics would need to rest on factual claims that would rationally motivate moral action.

Many utilitarians have followed the state-of-nature theorists in offering merely a morality for the lowest-common denominator—the sort that naturally emerges in response to troubling and apparently irresolvable cultural conflict. But as Perry notes, this move hides a presumption that ethics ought to gain "universal assent [and] *mathematically certain* and compelling answers, because in situations of crisis or deep-down pluralism relying on deliberation seems too weak." Rationality must be deductive, along the lines of geometry, a foundationalist image which Hobbes, Locke, and Bentham all suggest.[164]

But no substantive practical principles are either self-evident "relations of ideas," nor obvious "matters of fact" in the way utilitarians have hoped and sometimes argued. For instance, Singer has suggested that the universalizability condition for ethics makes utilitarianism the starting point beyond which any move needs to be justified: "The preference utilitarian position is a minimal one, a first base that we reach by universalizing self-interested decision-making."[165] As Perry shows, this move is highly doubtful. Hobbes and Locke, whose state-of-nature

[164] See Perry, "Where Did Utilitarianism Come From?," 31–32.
[165] Singer, *Practical Ethics*, 11.

suggestions Singer uses to support his argument, go far beyond preference universalization. And their sense that morality goes beyond that minimal condition bears the burden of proof no more than Singer's claim that it doesn't.[166]

The attempt to make the theory and practice of ethics acceptable to any rational being is one aspect of what MacIntyre calls Enlightenment liberalism. And MacIntyre notes that after centuries of systematic proposals and rigorous argument, none of the competing philosophical positions has been able to secure universal agreement about fundamental principles. This includes the failure to secure agreement about the nature of the final end of practical reason, or well-being.

MacIntyre sometimes suggests that the moralities of Enlightenment liberalism acknowledged no "overriding good."[167] Perry's critique of the Standard Story may apply to MacIntyre on this point; utilitarians do not *reject* but *reconceive* the overriding good. Specifically, they aimed to ground ethics in a widely attractive and empirical conception of well-being. This was a much thinner overriding good, but a good nevertheless.

MacIntyre does interpret later stages of liberalism as realizing the need for at least one common good—"the continued sustenance of the liberal social and political order,"[168] in which individuals may have radically different conceptions of the good life. And the difficulty of pursuing even this thin value forced liberalism, paradoxically, to become a tradition. Likewise, I propose that despite utilitarians' attempt to reduce the theory of well-being and transcend cultural particularity, even that minimal goal required the formation of a tradition. More on this shortly.

But there is a sense in which ethics, specifically the theory of well-being, may be empirical, or at least empirically accountable. Aristotelians take the study of human flourishing to be analogous to the study of the flourishing of any other life form. It is a matter of biological fact whether a human is functioning and thus flourishing.[169] And the relevant biological facts are accessible in part through empirical observation and ultimately the whole process of Aristotelian induction (*epagoge*). But the analogy is only partial in that human flourishing is much more

---

[166] Perry, "Where Did Utilitarianism Come From?," 31–32.
[167] MacIntyre, *WJWR*, 327.
[168] MacIntyre, *WJWR*, 345.
[169] Sober, *Philosophy of Biology*, 83ff; MacIntyre, *DRA*, chap. 7, esp. 63–67, 77–79; MacIntyre, *ECM*, 29–31.

complex, and thus harder to observe and evaluate, than plant and animal flourishing. Even for nonhuman animals, we must often rely on experienced "scientists" to make reliable judgments about flourishing. For Aristotelians, as well as utilitarians like Mill, the virtuous person is an expert in human flourishing. Importantly, the observation involved in the study of human flourishing is not empirical in a certain sense, because the relevant "sense-data" do not themselves deliver value judgments to observers. The value judgement is made in coordination with the observer's conception of the good.[170]

This is perhaps a step towards the best available compromise between Aristotelian and utilitarian conceptions of the epistemic accessibility of human flourishing. The blurring of the fact-value distinction which the Aristotelian standpoint requires may appear as a heavy cost for utilitarians, but that distinction in its strictest forms has manifested its own heavy costs in recent centuries. From the other direction, the utilitarian commitment to egalitarianism might suggest taking individuals' testimony about their own flourishing more seriously than an Aristotelian might be inclined to. Self-reports will be no more than, but also no less than, *prima facie* evidence.[171]

This discussion implies at least two contrasts between Aristotelian and utilitarian conceptions of moral inquiry: first, as to the *audience* of moral inquiry—*who is ethical inquiry for?*; second, as to the *decisiveness* of moral inquiry—*what can ethical inquiry conclude?*

The conception of ethics as an empirical science purports to make its conclusions acceptable to any rational individual. Naturally, the conclusions of ethics can be more decisive to the extent that the participants in the inquiry belong to a narrower pool. So the ability to *persuade widely* trades off with the ability to *conclude decisively*. Like Aristotle, Mill and Sidgwick acknowledge that the dispositions of students of ethics will shape ethics' ability to be practically beneficial for them.[172] In a sense, they restrict their audience to those who are "rational" in a sense that involves more than the logic of means and ends. It is a matter of

---

[170] Aristotelians or utilitarians might adopt what DiIulio describes as Mill's "value-sense theory" and "moral-sense theory."

[171] It may also need to be relatively clear to agents how their particular life can be or become *eudaimon*. Where Aristotelians insist that *eudaimonia* function as a target at which practical reason aims, utilitarians might insist that the plasticity of human nature requires that target to be relatively wide, or that there are multiple available targets.

[172] As I noted, Bentham must also assume the benevolence of his reformers, but it is not clear that he acknowledges this challenge.

contingent fact, which Aristotelians and utilitarians can dispute together, how many agents are in such a condition, but the Aristotelian tradition leans pessimistic.

The conception of ethics as an empirical science purports to make its conclusions practically decisive. Decisiveness is a standard more moderate than certainty, but it would still entail the ability to identify correct conduct despite any uncertainty. And this aim is reasonable—but only if the capacity to make a rational or moral choice is not construed as the application of a self-sufficient choice algorithm. The work of ethics involves applying a working value-conception to *my* unique circumstances, and in some sense only *I* can do this work. Nevertheless, Chapter II shows how a decision-making model might clarify my value-conceptions, even if they are quite complex.

Utilitarianism as a Tradition

Perry suggests that healthy ethical inquiry combines at least three modes or sub-inquiries: ethical theory, applied ethics, and the history of ethics.[173] He lauds utilitarianism for persistently developing *practical* aims alongside its *theoretical* project.[174] By contrast,

> what has been less common to the utilitarian tradition is its self-understanding *as a tradition*. It is a movement that is situated within a particular historical context, one that is extended over generations and achieves its coherence partly by reference to influential texts and decisive verdicts on practical questions.

MacIntyre explicitly names Bentham as a typical Enlightenment liberal, attempting to construct "a morality for tradition-free individuals."[175] Self-awareness about utilitarianism as a tradition seems to have grown from Bentham to Sidgwick, but then perhaps declined after utilitarianism's Victorian era. Mill was obviously self-aware about his own loyalty to his father's and Bentham's framework, and never gave it up while adjusting it to account for his own experiences, its external critics and internal challenges.[176] Sidgwick's comparative project and historical sensitivity approach Aristotle's own method, if not quite MacIntyre's.

Tradition awareness seems to have dissipated somewhat after Sidgwick. Scholars became specialized, devoting themselves to solving narrow problems under given presuppositions, without identifying within a tradition and attempting to defend a comprehensive system of thought and life. Nevertheless, we can find many strands of continuity between classical and

---

[173] Perry, "Introduction," 6.
[174] Perry, "Introduction," 6.
[175] MacIntyre, *WJWR*, 334, 353.
[176] DiIulio, *Completely Free*, 118–19.

post-classical utilitarianism, e.g., continuities from Sidgwick through Hare, Parfit, and Singer.[177] The degree to which these continuities have been recognized as such varies; some utilitarians more than others are self-consciously developing specific challenges that arise for a shared set of historically contingent commitments.

Perry names rabbinic Judaism, Aristotelianism, and Catholicism as clear examples of traditions, then proposes that

> [u]tilitarianism belongs on that list too, but has often eschewed that image, not recognizing that it too has 'authoritative' figures like Mill or Sidgwick or that it emerged as it did because it first faced a particular set of problems: in Bentham and Mill's case, the problems addressed by the great Victorian reform movements.[178]

Why have utilitarians often neglected the "traditional" mode of ethics? Perhaps because the utilitarian denial of its own status as a tradition

> appears to afford a strategic advantage. It allows utilitarianism to represent itself as the basic or default 'moral point of view' from which all other points of view begin, committed as it is to nothing more than happiness and fairness. Other moral theories add on to this basic background with various embellishments, such as virtue or God's command or categorical imperatives, but since these cannot garner universal assent, real ethics must be limited to what is morally basic. Unsurprisingly this basic ground turns out to be utilitarianism.[179]

This is just what I suggested with respect to Bentham's claim of incontestability for his principle of utility. It must be either a incontestable and vacuous tautology, or else a contestable and substantive thesis. But no tradition is basic; "each begins from somewhere," whether from the revelation of the Torah or the principle of human equality.[180]

Utilitarians may have several worries with the tradition-sensitive approach to ethics. One may be quickly assuaged: "[t]o ask that utilitarianism self consciously consider itself a tradition is not to stack the deck in favor of religious ethics." Theologians do use the word 'tradition' more than philosophers, and "[i]t is one of theology's gifts to the wider academy—whatever else one thinks of religion or theism or Christianity—that it models how to integrate systematic reflection on a body of knowledge with the study of texts, together with history, together with practical ethics."[181] But tradition is not the exclusive domain of religion. In fact, utilitarianism's

---

[177] Perry, "Introduction," 2–3.
[178] Perry, "Introduction," 7.
[179] Perry, "Introduction," 7.
[180] Perry, "Introduction," 7.
[181] Perry, "Where Did Utilitarianism Come From?," 17.

history as a distinct movement, especially the project of progressive social reform from Bentham through Singer, is one of its attractions. And as we have already seen, the utilitarian-Christian rivalry is thoroughly misleading.[182]

A deeper worry regards the inescapability of a contingent starting point in ethical reasoning. MacIntyre has argued that this inescapability need not preclude the possibility of rational objectivity. I cannot here say enough to satisfy utilitarians on this point, but I refer them again to the work of MacIntyre and his commentators on tradition, truth, and justification.[183] What the position does certainly entail is that moral understanding requires attention to historical context along with theoretical analysis, attention to how a tradition "conceives itself as a conversation over time about how to live morally." Mill's own conception of utilitarian morality as a "Nautical Almanack," to be revised through experience over long histories, might reach toward this awareness.[184]

Both these concerns stem from utilitarianism's Enlightenment origins, which rejected both particular religious traditions and commitments, and tradition and commitments as such, without always clearly distinguishing the two. Thinkers like Locke, Kant, and Mill "did not see their view as one plausible option among many, but instead each was equally convinced that his theory alone was the one that any rational person would accept." But early modern philosophers characteristically didn't think historically or genealogically: "Each thinker interacted with his immediate predecessors as foils, but almost never did they step outside their immediate context to ask how they got there, or what conceptual burdens they may have brought."[185] Enlightenment thinkers referred to their historical inheritance primarily to disparage it, reinforcing their sense of having transcended their context and eliminated presuppositions.[186]

---

[182] Perry also points out that Singer's Shallow Pond parable bears more than passing resemblance to the Good Samaritan, and "utilitarians were the only allies Christians had in demanding radical, sacrificial generosity and giving to the poor; the classical liberals thought it might be nice to give, but not obligatory." Perry, "Introduction," 7–8; Perry, "Where Did Utilitarianism Come From?," 16.

[183] Lutz, *Tradition in the Ethics of Alasdair MacIntyre*; MacIntyre, *WJWR*, esp. Chs. I, XVII-XIX; MacIntyre, *3RV*, esp. Chs. VI-IX; MacIntyre, *ECM*.

[184] Mill, *Utilitarianism*, II; DiIulio, *Completely Free*, 118–19.

[185] Perry, "Where Did Utilitarianism Come From?," 17.

[186] Perry admits these critiques are less apt of the historically-sensitive Sidgwick, who "knew that he could not do his job well without thoroughly understanding all of his predecessors—Hutcheson, Clark, Shaftesbury, Hume." Sidgwick also backed off the ambition of universal rational persuasion, lowering his epistemic standard to persuasion of a rigorously-defined "rational person." Despite Sidgwick's attempt to bring the utilitarian tradition to some degree of self-awareness, it seems to have been dropped by subsequent utilitarians.

But no theorist draws on abstract rationality alone. Utilitarianism in particular "did not emerge promiscuously and spontaneously from a wide range of human cultures (as, for example, the Golden Rule did), but only in a very precise historical context and via a particular intellectual genealogy."[187] MacIntyre often complained about the treatment of "ethics" as a single subject about which philosophers through history simply make different claims—in fact, he thinks, the history of ideas and analysis of ideas cannot be sharply distinguished, so to abstract thinkers from their context commits us to misunderstanding them.[188]

The utilitarian insistence on a rational ethics follows from a widespread early modern contrast between matters of aesthetic taste and matters of rational judgment. Despite their empiricist affinities, utilitarians tend to follow Kant in Kant's disagreement with Hume.[189] Hume, along with intuitionists, realized that moral science couldn't be like mathematics, because it derives from human passions. But both Bentham and Kant thought this rendered ethics too contingent and unsystematic. Since morality couldn't be founded on the passions, it must be founded on reason.[190]

This dilemma between morality as taste and morality as mathematics fails; "sound moral judgement might, in some respects, be analogous to taste."[191] Judgments of taste may be, as Aristotle understood them, judgments about excellence for some purpose, intelligible in the context of a practice or functional concept. In the domains of activity MacIntyre calls 'practices', wide disagreements are possible but cast no doubt on the objectivity of the criteria: "such judgements will not have the precision of a closed theoretical system like geometry, but why should they?"[192] Even utilitarians have to concede that

---

[187] It is instructive to contrast Perry here with Geoffrey Scarre, who happily casts a wider net for utilitarians and "proto-utilitarians," a category to which Epicurus, Mozi, and Jesus all belong. I take it these two ultimately disagree about little of substance, but they use the word 'utilitarianism' quite differently. For Perry, it refers *ostentively* to a particular social and intellectual movement; for Scarre, it refers *descriptively* to a set of doctrinal commitments, construed broadly enough to include thinkers from independent traditions.Perry, "Where Did Utilitarianism Come From?," 16; Scarre, *Utilitarianism*.

[188] MacIntyre, *AV*, 11; MacIntyre, *SHE*, 1–4. T. H. Irwin and J. B. Schneewind are, I think, MacIntyre's most fruitful critics on this point. See Irwin, "Tradition and Reason in the History of Ethics"; Schneewind, "Moral Crisis and the History of Ethics."

[189] See, e.g, Singer, "Ethics and Intuitions."

[190] Perry, "Where Did Utilitarianism Come From?," 33; MacIntyre, *AV*, 49. Sidgwick believed there was such a thing as objectively good taste: "it is only the judgement of persons of good taste that we recognize as valid in respect to the real goodness of the thing enjoyed." But he still supposed that since "many pleasures [are] available from any given object or practice or conduct," taste couldn't give morality its "authoritative prescription." Henry Sidgwick, *The Methods of Ethics*, 7th ed., IX.2, 108ff.

[191] Perry, "Where Did Utilitarianism Come From?," 33.

[192] Perry, "Where Did Utilitarianism Come From?," 36.

> it is very often hard to see what actions will [maximize utility], especially in the very long run. If so, our judgements about right conduct may be more like judgements about, say, the best art, or horse or wine, and less like a mathematical proof.[193]

Once it is admitted that moral inquiry cannot ensure the decisiveness of its conclusions, utilitarians might have little reason to insist on thin value conceptions, little reason to reject the plurality of value or the criticism of human dispositions.[194]

Perry offers Jonathan Haidt's work as a scientific approach to ethics that nevertheless recognizes its analogy to taste. Haidt sees normative concepts emerging in human cultures and individuals from several "moral foundations." These concepts are partially innate, but we acquire them primarily through socialization. Just as a science of human taste requires study of the tongue, the science of ethics requires study of "moral taste receptors."[195] Since tastes diverge so widely, the analogy may seem to suggest that moral truths are radically contingent or relative. Preferentist utilitarians preclude relativity by saying the moral goal is simply to maximally satisfy the tastes of each.[196] Haidt instead suggests that empirical study can produce substantive results about moral tastes. Admittedly, a moral theory cannot be read straight off empirical findings—theorists must work out how to turn the moral foundations into a coherent conception of "the happy human life and in the flourishing society."[197] They will subject moral to criticism along the way.[198]

Utilitarians like Singer may find Haidt still too teleological—why think our intuitions tell us anything at all of value? Haidt, like Aquinas, might be accused of violating "no ought from is." Perry objects:

> understood in those terms, does not all of utilitarianism break the rule? Is not the whole point of consequentialism that the 'is' of making-happy entails the 'ought' of pursuing it? …Singer needs to show why desiring fruit is morally relevant, Haidt needs to show that intuitions are morally relevant, and Aquinas needs to show why biological capacities are morally relevant."[199]

---

[193] Perry, "Where Did Utilitarianism Come From?," 37.

[194] Perry offers Peter Singer Jonathan Haidt's "richer conception of moral goods," which is semi-utilitarian with Humean and Durkheimian influence. He thinks Haidt "might provide something of a preliminary point of contact between various forms of eudaimonism…and Singer's utilitarian alternative." Perry, "Where Did Utilitarianism Come From?," 37.

[195] Haidt, *The Righteous Mind: ?*, 57.

[196] Perry, "Where Did Utilitarianism Come From?," 41.

[197] Perry, "Where Did Utilitarianism Come From?," 42.

[198] Perry, "Where Did Utilitarianism Come From?," 42.

[199] Perry, "Where Did Utilitarianism Come From?," 44.

Of course, a fact *by itself* isn't morally relevant; our observations become distinctly moral when we subject intuitions, sentiments, and desires to critical reflection. This is the function of reason, Haidt's rider on the elephant. While we cannot just read a moral theory straight off Haidt-style research, we can reason about how to integrate the various "foundations" in a coherent conception of "the happy human life and in the flourishing society."[200]

The Challenges of Value Plurality and Criticism

In Chapter 2, I discuss the divergence of characteristic Aristotelian and utilitarian attitudes towards the plurality of values. I argue for a moderate position—irreducible but commensurable plurality. Practical commensurations among values can be rationally justified, but they will be contestable, emerging not from an algorithm but from individual experience and within social contexts like an Aristotelian *polis* or MacIntyrean tradition. But utilitarian resources have an important function in the picture. The formal machinery of rational choice theory is in key respects designed explicitly to *clarify* the best way of making comprehensive evaluative judgments which account for all relevant values.

In Chapter 3, I discuss the divergence of characteristic Aristotelian and utilitarian attitudes towards the criticism of human tastes and desires. I argue for a moderate position—a nontrivial but bridgeable gap between what seems good to a person and what is actually good for them. The supposition of such a gap is *nontrivial* in that it relies on contestable claims about human nature, not merely logical principles or empirical facts. The gap must be *bridgeable* in the sense that agents are not alienated from their own flourishing. On this picture, the achievement of flourishing requires an alignment of means to ends *and* an alignment of present ends to proper ends.

Both these issues embody the broader epistemic or meta-ethical issues at stake between the two traditions. In fact, both may be considered watersheds between utilitarianism and Aristotelianism or eudaimonism.[201] So naturally, some utilitarians will remain dissatisfied with

---

[200] Perry, "Where Did Utilitarianism Come From?," 42.

[201] Kelly, "MacIntyre's Critique of Utilitarianism," 136. Perry suggests these features are apt of classical utilitarianism as distinctly shaped by Bentham and Mill, but do not characterize the whole past of the utilitarian tradition, nor its necessary future. Perry nevertheless labels himself a eudaimonist but not a utilitarian. He considers ethics to be fundamentally about "flourishing or well-being or happiness," including God and nature in in the moral sphere alongside humans. Perry allows this view to be categorized with utilitarianism as both teleological and consequentialist in some sense. He notes that although the original Anscombian sense of 'consequentialism' did not include eudaimonism, it now has a broader meaning which perhaps includes "any theory of ethics that is, at bottom,

the view of moral inquiry on offer here. For Singer, "one of the great advantages of any form of utilitarianism that is based on just one value… is that there is no need to explain how different values are to be traded off against one another."[202] But this is not an argument in favor of monism *per se*, only the suggestion that value monism would be theoretically convenient for philosophers whose "ideal form of moral reasoning has the precision of mathematical deduction." In other domains, such as aesthetics, we deliberate, evaluate, and judge without this expectation.[203]

I can only offer my compromise positions to utilitarians and Aristotelians hoping they will be recognized in substantial continuity with each tradition's core historical commitments.

V: Conclusion: Utilitotelianism? Aristotilitarianism?

A central aim of my project is to indicate how utilitarians may transcend their own internal conflicts through self-awareness about their history and development, particularly in relation to Aristotelian ethics. The recognition needed is one that, so far, the tradition has only partially achieved. Utilitarians *have partially* achieved it insofar as some utilitarians repudiate some of the mistaken extremes of the Enlightenment project of which utilitarianism was originally a part. Foremost among Enlightenment mistakes were the obscuring of tradition-bound inquiry and the subsequent imposition of unachievable epistemic standards for the project of ethical inquiry. Utilitarians have *only partially* achieved the needed recognition insofar as those epistemic standards endure, whether implicitly or explicitly.

Given this recognition and given the two traditions' core commitments and challenges, a mutually enriching synthesis between Aristotelianism and utilitarianism is both possible and desirable to a greater extent than has yet been acknowledged by members of either tradition. I have identified and addressed several key points of interaction—some in tension and some in

---

concerned with promoting well-being or flourishing or happiness." Perry, "Where Did Utilitarianism Come From?," 18–19.

[202] Singer, *Practical Ethics*, 169.

[203] Perry continues:

> What is the best painting? This cannot be answered on monist value grounds, for all sorts of factors might be relevant: the painting's age, its continuity or contrast to previous works, its place within a particular school of painting, its meaning within the artist's own life, its physical condition, its sociopolitical significance, and so on and so forth. Now, choosing the best painting in this way does indeed 'sacrifice one of the great advantages of' choosing on a single value. How much simpler it would be to assert a single value, like age or price, to make all aesthetic values commensurable! But aesthetic judgement is not the sort of deliberation for which mathematical precision is appropriate.

Perry, "Where Did Utilitarianism Come From?," 48.

harmony—between the Aristotelian and utilitarian traditions. The extent to which the distinctives of each tradition can be held together must be evaluated case-by-case. In principle, any case could reveal a fundamental incompatibility, which would function as a dividing line between faithful members of each tradition. From where I stand, none of the points on which particular Aristotelians and utilitarians have genuinely contradicted each other are essential to Aristotelianism or utilitarianism. On at least the key points treated in my project, each tradition can accommodate its rival's insights without internal contradiction.

2. Irreducible but Commensurable: Moderating Aristotelians and Utilitarians on Value Conflicts

Introduction

We humans sometimes confront difficult value comparisons—occasions on which our choice options seem good, or good-for-us, in quite distinct ways. These occasions are frequent and significant enough that any robust account of practical rationality and well-being must address them. Unsurprisingly, then, the problem of how to understand and respond to difficult value comparisons has great significance in the rivalry between the Aristotelian and utilitarian traditions.

In my view, the Aristotelian tradition has the conceptual resources to rightly respect the plausible claim that not all good things are good in the same way or because they share the same feature. Call this feature (and principle) the *irreducibility* of plural values. On the other hand, the utilitarian tradition has the conceptual resources to rightly respect the plausible claim that rational decision-making requires some non-arbitrary procedure for comparing choice options, which in turn requires that values are somehow commensurable. Call this the *commensurability* of plural values. In my view, not only can value irreducibility and commensurability be held together consistently; plausibly, utilitarians and Aristotelians both need to hold them together in their accounts of practical reasoning.

Unfortunately, the Aristotelian case for value irreducibility has often gone so far as to make values incommensurable, and the utilitarian case for value commensurability has often gone so far as to make values reducible. But in reckoning with the strengths and limitations of both kinds of arguments, we can develop a robust account of decision-making despite value conflicts that makes use of both traditions' conceptual resources and respects both traditions' core commitments. On this account, values are irreducible but commensurable. Value commensuration is a cognitive process agents must actively undertake, and the evaluations it involves can be more or less justified, the choice in which it issues mor or less rational. Commensuration is neither automatic nor easy, requiring agents to keep all their relevant values salient as they integrate them into comprehensive evaluations of outcomes. The capacity to make these evaluations partially constitutes personal both practical rationality and psychological integrity. It is a capacity which may be more or less developed in an agent, and its development is shaped both by agents' cultural environment and their engagement with that environment.

In what follows, I begin to substantiate these high-level claims through an investigation of a few particular perspectives on the topic. MacIntyre, naturally, is my representative Aristotelian, both in his extreme incommensurabilism in *After Virtue* (Part I) and in his more moderate view in later work (Part IV). After briefly noting the extreme commensurabilism of Bentham's classical utilitarianism (Part II), I take Stephen Ellis as developing utilitarian resources toward a more moderate commensurabilism that allows irreducibility (Part III). Out of these sources and their interactions I develop a constructive proposal for mutual inter-tradition enrichment on the problem of difficult value comparisons, avoiding both reductivism and incommensurability. Ellis's defense and explication of value commensuration can find a home in MacIntyre's Aristotelian normative framework, and some formal work in decision theory can precisify that framework. Moreover, MacIntyre's narrative history of modernity contextualizes the problem of incommensurability as a contingent phenomenon characteristic of our time and place.

I: Extreme Incommensurabilism (*After Virtue*)

MacIntyre has long been fond of 'incommensurable' and related terms, using them relatively often and not always in precisely the same sense. It is therefore important at the outset of my discussion to carve away MacIntyre's uses of 'incommensurability' in connection with rival moral traditions and the fragmentary moral concepts in which contemporary moral debate is conducted.[204] I do suspect it is not mere homonymy that connects incommensurability among concepts and traditions and incommensurability among values, goods or ends. But I need not explore this connection here. Now, as to conflicts between rival goods, how does MacIntyre understand them in *After Virtue* (hereafter *AV*)?

Three Interpretations of Tragic Conflict

MacIntyre's polarizing 1981 classic is, of course, an attempt to vindicate a broadly Aristotelian conception of ethics against the Enlightenment project and its Nietzschean consequences. In discussing value conflicts, he contrasts Aristotle's view with a "sharply

[204] MacIntyre himself acknowledges the distinction when he identifies Cardinal Pole and Marquis of Montrose as moral exemplars, whose virtues enabled them to act well "even in situations defined by the necessity of tragic, dilemmatic choice." He says that "[s]uch choices, understood in the context of the tradition of the virtues, are very different from those which face the modern adherents of rival and incommensurable moral premises." MacIntyre, *AV*, 223.

contrasting modern tradition," which he associates with Max Weber and Isaiah Berlin.[205] On this view,

> the variety and heterogeneity of human goods is such that their pursuit cannot be reconciled in any single moral order and that consequently any social order which *either* attempts such a reconciliation *or* which enforces the hegemony of one set of goods over all others is bound to turn into a straitjacket and very probably a totalitarian straitjacket for the human condition.[206]

Further, this "modern" view implies that the choices between goods, and presumably the choices between virtues, cannot even be expressed as judgments with truth values.

MacIntyre of course rejects this view, but not by offering Aristotle's as decisively preferable. In an acutely uncharacteristic moment, MacIntyre shows explicit sympathy for a particular modern liberal critique of Aristotle, namely, that he "offers too simple and too unified a view of the complexities of human good." The city of Athens, not to mention the wider ancient world, involved "a recognition of a diversity of values, of conflicts between goods, of the virtues not forming a simple, coherent, hierarchical unity." Aristotle allegedly idealizes the moral life and "exaggerate[s] moral coherence and unity." Accordingly, his commitment to the unity of the virtues seems to go beyond what he is able to establish about the particular "interrelationships between different virtues and vices."[207]

Without granting the Weberian claims, MacIntyre takes these challenges seriously and suggests we ask what motivates the strength of Aristotle's commitments here. Plato turns out to be partially to blame, since Aristotle inherits the unity of the virtues "directly from Plato," as part of "an hostility to and denial of conflict either within the life of the individual good man or in that of the good city." Plato was convinced that "[t]here cannot be rival goods at war with each other." He allowed disagreements between people about the virtues, but not disagreements between the virtues themselves. If two virtues seem at odds, then some party must be mistaken, and "we can rationally settle all such disputes and arrive at a single rationally justifiable account of and list of the virtues." The virtues cannot even be in tension on particular occasions or temporarily. It cannot be that, e.g., Philoctetus' duties as Odysseus' friend are genuinely at odds

---

[205] MacIntyre perhaps had in mind Berlin, *Four Essays on Liberty*, 168; Weber, "Science as a Vocation," 22.

[206] MacIntyre, *AV*, 142.

[207] MacIntyre, *AV*, 157.

with his duties to treat Neoptolemus honorably. In fact, Plato makes the even stronger claim that the virtues are not only necessarily compatible but also necessarily co-present.

MacIntyre notes that Aristotle and Aquinas both follow Plato in presupposing the existence of "a cosmic order which dictates the place of each virtue in a total harmonious scheme of human life. Truth in the moral sphere consists in the conformity of moral judgment to the order of this scheme."[208] This is why for Plato and then Aristotle, conflict is an evil, to be counteracted by the harmony of the virtues in the individual and the state; "the good life for man is itself single and unitary, compounded of a hierarchy of goods. It follows that conflict is simply the result either of flaws of character in individuals or of unintelligent political arrangements."[209]

Aristotle's commitment against conflict leads the *agon*—the athletic contest—to be "displaced from its Homeric centrality" across his work.[210] In Aristotle's epistemology, "dialectic is no longer the road to truth, but for the most part only a semi-formal procedure ancillary to enquiry." Whereas Socrates dialogued with individuals and Plato wrote dialogical works, Aristotle "produces expository lectures and treatises."[211] Aristotle's theology contrasts with the playwrights' along the same lines:

> for Aristotle that particular appeal to the divine which in both Aeschylus and Sophocles signals the recognition of tragic impasse can have made no realistic sense. The impersonal unchanging divinity of which Aristotle speaks, the metaphysical contemplation of which furnishes man with his specific and ultimate telos, can itself take no interest in the merely human, let alone in the dilemmatic.[212]

Most relevantly, Aristotle's aversion to "the dilemmatic" forces him to adopt a particular poetic theory, in which

> the tragic form of narrative is enacted when and only when we have a hero with a flaw, a flaw in practical intelligence which springs from inadequate possession or exercise of some virtue. In a world in which everyone is good enough therefore there would be no tragic hero to be portrayed.[213]

But this enormously influential interpretation of tragedy is a *mis*interpretation of the genre, which explores exactly what Plato took to be impossible: irresolvable conflicts of duties,

---

[208] MacIntyre, *AV*, 142.
[209] MacIntyre, *AV*, 157.
[210] MacIntyre, *AV*, 157.
[211] MacIntyre's contrast here should be qualified by noting that Aristotle's lost works include those dialogues of which Cicero spoke so glowingly. Shields, *Aristotle*, 25. Nevertheless, the point stands that Plato did not produce volumes of prose discourse like Aristotle's.
[212] MacIntyre, *AV*, 157–58.
[213] MacIntyre, *AV*, 163.

loyalties, and "allegiances to incompatible goods." The inadequacy of Aristotle's moral psychology becomes inadequacy in reading dramas like *Oedipus Rex*:

> the conflicts of tragedy certainly may in part take the form that they do because of the flaws in Antigone and Creon, Odysseus and Philoctetes; but what constitutes those individuals' tragic opposition and conflict is the conflict of good with good embodied in their encounter prior to and independent of any individual characteristics.[214]

Homer and Sophocles acknowledged what Aristotle would not: "that tragic conflict is the essential human condition—the tragic hero on Aristotle's view fails because of his own flaw, not because the human situation is sometimes irremediably tragic."[215]

MacIntyre contrasts Sophocles' exemplary view with both the Platonic and Weberian theses:

> There are indeed crucial conflicts in which different virtues appear as making rival and incompatible claims upon us. But our situation is tragic in that we have to recognize the authority of both claims. There *is* an objective moral order, but our perceptions of it are such that we cannot bring rival moral truths into complete harmony with each other and yet the acknowledgment of the moral order and of moral truth makes the kind of choice which a Weber or a Berlin urges upon us out of the question. For to choose does not exempt me from the authority of the claim which I chose to go against.

This is why Sophoclean tragedies conclude with the *deus ex machina*, the pronouncement of an authoritative verdict by a god, which "always *ends* rather than *resolves* the conflict."[216] Divine intervention does not preclude human responsibility; it highlights both the reality of the moral order and its occasional inaccessibility to human understanding. The self "remains accountable… for the way in which it handles itself in those conflicts… it can indeed win or lose, save itself or go to moral destruction." The divine order grounds evaluative judgments by impelling our pursuit of certain ends.[217]

MacIntyre sides with Sophocles, identifying Aristotle's commitment to the unity of the virtues as one of the central liabilities which will have to be overcome in resuscitating Aristotelianism for modernity. In his own voice, he asserts that humans must in fact decide between "rival and contingently incompatible goods." MacIntyre's qualifier "contingently" here would seem to imply a weaker claim than true incommensurability, as if plural goods do not

---

[214] MacIntyre, *AV*, 164–65.
[215] MacIntyre, *AV*, 157.
[216] MacIntyre, *AV*, 143.
[217] MacIntyre, *AV*, 145.

exclude each other by their nature. But the stronger claim is clear in the following key passage, in which he says of tragic situations that

> *both* of the alternative courses of action which confront the individual have to be recognized as leading to some authentic and substantial good. By choosing one I do nothing to diminish or derogate from the claim upon me of the other; and therefore, whatever I do, I shall have left undone what I ought to have done.[218]

MacIntyre contrasts this choice with the contemporary choice of allegiance to one moral principle over another. The two kinds of choices involve different conceptions of obligation. Modern moral principles, whether deontological or utilitarian, are generally understood to conform to Kant's Law, according to which "ought implies can." By contrast, the tragic agent cannot do all that they 'ought' to do; they are obligated beyond their capacity.[219]

Although MacIntyre favors Sophocles against Aristotle on tragic conflict, he still aims to maintain a broadly Aristotelian outlook. Regarding the centrality of the *polis*, he affirms John Anderson's "Sophoclean insight" that social institutions are oriented not around static ends but around dynamic conflicts over ends; these conflicts are a key domain for the acquisition and exercise of the virtues.[220] MacIntyre concludes his chapter on Aristotle with the claim that "it is through conflict and sometimes only through conflict that we learn what our ends and purposes are."[221]

As an Aristotelian, MacIntyre must also maintain the possibility of agents' responding to the reality of tragic conflict in a practically rational. How can he? In introducing this question, MacIntyre notes a view he attributes to J. L. Austin, according to which

> *either* we can admit the existence of rival and contingently incompatible goods which make incompatible claims to our practical allegiance *or* we can believe in some determinate conception of *the* good life for man, but that these are mutually exclusive alternatives.[222]

MacIntyre asserts to the contrary that we cannot rule out that the good life for man might precisely require successfully navigating "the tragic confrontation of good with good." He criticizes Austin for assuming without argument that "contingently incompatible" goods would

---

[218] MacIntyre, *AV*, 223–24.

[219] MacIntyre, *AV*, 224.

[220] For an argument that MacIntyre exaggerated the extent to which Aristotle exaggerated the unity of the *polis*, see Yack, "Community and Conflict in Aristotle's Political Philosophy."

[221] MacIntyre cites Anderson, *Studies in Empirical Philosophy*, xxii; MacIntyre, *AV*, 165.

[222] MacIntyre, *AV*, 223. MacIntyre probably has in mind Austin, "A Plea for Excuses: The Presidential Address," 29n16.

rule out the possibility of better and worse responses to tragedy.[223] Despite the absence of a *right* choice, "the moral task of the tragic protagonist may be performed better or worse, independently of the choice between alternatives that he or she makes." For one thing, they may proceed through the tragedy

> heroically or unheroically, generously or ungenerously, gracefully or gracelessly, prudently or imprudently. To perform his or her task better rather than worse will be to do both what is better for him or her *qua* individual and *qua* parent or child or *qua* citizen or member of a profession, or perhaps qua some or all of these.[224]

Tragic dilemmas are not counterexamples to the idea that claims like the following have objective truth values: "To do *this* in *that* way would be better for me and/or my family, city, or profession." Such claims about what is good for me are justifiable and intelligible in the context of my life as unified whole. In MacIntyre's view, it is because modernity lacks such a conception of a life's unity that it cannot consider moral judgments, including ascriptions of virtues and vices, to be matters of fact.[225]

In sum, this MacIntyre accepts the possibility of genuine, rather than merely apparent, conflict between goods to be practically pursued by Aristotelian agents. He thinks certain goods make ineliminable claims upon agents, such that when they are available to the agent but mutually exclusive, whichever good is neglected still, in some sense, "ought" to have been pursued. To deny the impasse is to paper over the legitimate claims of some "authentic and substantial" value. Nevertheless, MacIntyre maintains that we can make better or worse decisions between rival goods, where these decisions are evaluable on the basis of the integrity of an individual's whole life. In my concluding analysis, I comment on the strengths and limitations of MacIntyre's characteristically provocative discussion.

II: Extreme Commensurabilism (Hedonistic Utilitarianism)

While utilitarianism is in crucial respects a product of enlightenment liberalism, not everything MacIntyre says about "liberalism" and "modernity" is true of utilitarianism. For example, Bentham is quite clearly committed to the concept of an objective overriding good—maximal pleasure. Admittedly, compared to an Aristotelian account of the good, Bentham's is thin, reductive, and subjective in the sense of being a private mental state. And Bentham does

---

[223] MacIntyre, *AV*, 223–24; MacIntyre probably has in mind Austin, "A Plea for Excuses: The Presidential Address," 29–30.

[224] MacIntyre, *AV*, 224.

[225] MacIntyre, *AV*, 224–25.

acknowledge that a heterogenous set of experiences is necessary to achieve this good. But on the other hand, Bentham thinks that all experiences can ultimately be commensurated and evaluated with respect to the quantity of pleasure they contribute to the total.

For Bentham, the felicific calculus enables normative judgments to be a matter of evaluative fact rather than merely emotive or imperative expression. A reductive conception of value was necessary to set ethics on its new scientific foundation and eliminate the arbitrariness of emotion and prejudice. To explain and justify normative rational and moral standards, actions must be evaluated according to the outcomes they produce, and the relative goodness of possible outcomes must be a matter of scientific fact. Such facts must be empirically accessible—specifically, they must refer to a measurable quantity. Despite the plurality of particular human aims, these aims are aimworthy because they instantiate some single common feature. The apparent diversity of goods is really a diversity of sources or causes of the one good. So, the plurality of pleasure sources only matters because of contingent considerations, like diminishing returns from, or limited access to, particular pleasure sources.

Reductive hedonism faces deep challenges. I discuss some of these at more length in chapter 3. Here, I mean only to suggest that Bentham's hedonism fails on its own terms; that is, it fails as a fully reductive (quantitative) conception of value. Suppose we grant the central claims that pleasure is the only intrinsic good, all goods being reducible to it, and pain the only intrinsic bad, all bads being reducible to it. Even so, pleasure needs to be commensurated against pain, because these are not obviously just two valences of the same fundamental experience, representable (e.g.) by positive and negative points on the y-axis.[226] Likewise, duration of hedonic experience needs to be commensurated against intensity of experience, because (again) these are not obviously just two dimensions of the same thing, representable by two dimensions (length and height) on the xy-plane.[227] To meet Bentham's standards, these commensurations would require not just any weights or proportions, but a formula that describes pain and pleasure, and duration and intensity, in terms of a single quantitative currency. And this seems out of reach—there seem to be qualitative differences between pain and pleasure as such, and between duration and intensity of hedonic experience.

---

[226] Watson, *Mood and Temperament*, 26–33, 44–54; Katz, "Pleasure," Note 1; Bradley, "Objective Theories of Well-Being," 230.

[227] Donner, "Mill's Theory of Value," 121–22; Kahneman, *Thinking, Fast and Slow*, 377–85; "Introduction," 10–11.

In other words, certain experiences seem irreducible. We need not abandon, as MacIntyre does, the initial utilitarian aspiration to a non-arbitrary way of comparing the overall value of outcomes. But we perhaps should abandon the construal of "overall value" as a reductive aggregation of a single quantity. We need an all-things-*considered* evaluation, rather than an all-things *collapsed* evaluation, and this is just what Ellis proposes.

III: Moderate Commensurabilism (Ellis)

From among the great deal of recent literature on difficult value comparisons,[228] I want to draw three central insights from Stephen Ellis's work.[229] Firstly, the main argument that some values are incommensurable rests on a conception of human motivation according to which "each person's basic ends are not only irreducible but also incompatible with one another."[230] Secondly, this conception and the argument it underlies is faulty, because any particular value or basic desire merely

> underwrites a *ceteris paribus* evaluation. Such assessments are mutually compatible because the observation that there is something to be said for an outcome from a particular perspective allows for any ultimate evaluation of that outcome. Values can be irreducible without thereby being incommensurable.[231]

Finally, Ellis also proposes a process of "experience-guided calibration" by which individuals can learn, over time, to commensurate in a way that respects their value system as a whole and thus preserves their personal integrity.[232]

What exactly would it mean for values to be incommensurable? Begin by noting that on at least some occasions, we seem to compare outcomes, or choice options, with respect to some particular end, which Ruth Chang calls a *covering value*.[233] I generally compare gas stations by considering which has the lowest price per gallon—fuel cost is the covering value with respect to which I compare my options. So value comparisons are three-place relations: "*x* is better than (worse than, etc.) *y* with respect to V."[234] 7-Eleven is better than Conoco with respect to fuel

---

[228] Nien-he Hsieh and Andersson, "Incommensurable Values"; Chang, *Incommensurability, Incomparability, and Practical Reason*; Andersson and Herlitz, *Value Incommensurability*.

[229] Ellis, "The Main Argument for Value Incommensurability"; Ellis, "Multiple Objectives."

[230] Ellis, "The Main Argument for Value Incommensurability," 27. Ellis credits Davidson with this view. Davidson, "How Is Weakness of Will Possible?"

[231] Ellis, "The Main Argument for Value Incommensurability," 27. Ellis credits Davidson with this view. Davidson, "How Is Weakness of Will Possible?"

[232] Ellis, "Multiple Objectives," 330.

[233] Chang, "Introduction," 5. Chang's work has had a central role in shaping contemporary work on value conflicts.

[234] Ellis, "The Main Argument for Value Incommensurability," 29.

cost. Note also that some values are composite, being comprised of multiple covering values. I generally compare food choices by considering cost, taste, and healthiness. Chang calls these *contributory values*.[235]

The concept of incommensurability can thus be framed in Chang's terms like this: apparent incommensurabilities arise when different contributory values favor different actions, and those values cannot be compared with respect to the covering value.[236] A milkshake is tastier but less healthy than a protein shake, and it's not clear in what proportion the two considerations contribute to the value of snacking. If health and gustatory pleasure are *incommensurable*, the outcomes that realize those values *incomparable*—neither better than, nor worse than, nor equal to each other.[237] If rationality requires choosing an option at least as good as any other,[238] then incommensurability would seem to imply that there's no rational way to choose my snack. What considerations support this kind of claim?[239]

Value incommensurability has been asserted on various grounds, but Ellis suggests they follow a general pattern:

> Starting with the observation that people have a number of basic goals, incommensurabilists argue that these ends are not only irreducible but also that their assessments are incompatible with one another in an important way."[240]

---

[235] Chang, "Introduction," 7.

[236] Ellis, "The Main Argument for Value Incommensurability," 29.

[237] One way of addressing difficult value comparisons is to propose a fourth value relation, such as "parity" (Chang) or "rough equality" (Parfit) besides the standard three relations. Arguably, such accounts allow choice options, and perhaps even values, to be *comparable* without being *commensurable*. I suspect that when agents deliberate in the face of uncommensurated values, they will still have to treat their options as conforming to the traditional trilemma—better, worse, or equal—and choose an option at least as good as any other. The vocabulary of a fourth value relation may help us avoid awkwardly precise claims about how radically different values compare, but I doubt it will function significantly in practical reasoning.

[238] For references to the literature on this point, see "Introduction."

[239] Andersson and Herlitz seem to think this is the wrong question:

For those who accept that there is a plethora of practical reasons and values and adhere to some pluralistic account of reasons and value, it seems highly plausible that these can come into conflict in the sense that we cannot determinately judge how they relate to one another. At the very least, one needs an argument for the opposite view, the assumption that it is always, in principle, possible to determine which of two options is better than the other or that they are equally good. The burden of proof seems to lie with those who reject incommensurability. "Introduction," 7.

On the contrary, given that apparently rational agents make their way through difficult value comparisons *somehow*, the burden of proof would seem to be on those who allege that whatever psychological process this involves is an *irrational* one.

[240] Ellis, "The Main Argument for Value Incommensurability," 29–30.

The beginning premise of the incommensurability argument is value pluralism, according to which multiple goods, ends, or basic desires are irreducible to one another.[241] Each value makes an independent claim on the agent, precluding any shared "currency" of comparison. Perhaps what makes health good for me is just not the same thing that makes pleasure good for me. And this is perhaps why many choices between such values involve a felt remainder of regret for any neglected values.[242]

But irreducibility alone does not entail incommensurability. As Sen says, "[t]wo things may be diverse and may not be measurable in any common unit, and yet very easy to order in line with what we value."[243] For example, we can and do rank various flavors despite their vast variety.[244] What I like about Pad See Ew is irreducibly unlike what I like about chocolate cake, but I decisively prefer Pad See Ew. Likewise, values may be multiple, independent, and irreducible without being essentially competitive, incomparable, or incommensurable. So, this inference must be justified on some further grounds.

Many incommensurabilists note the intense difficulty of value commensuration. They argue that agents are in positions where they ought to be able to "see" value weights if they existed to be seen. So, the fact that we cannot thus see value weights evidences their non-existence.[245] But as Ellis argues, there are better ways to explain the difficulty of value comparisons than by denying the existence of value weights. He notes first that ranking options may simply be unlike sensory perception—values do not have weights "attached" for us to just notice and then account for in deliberation. Comparing ends might instead be difficult because we lack sufficient self-knowledge, of the kind that Ebenezer Scrooge acquires in his spiritual visions. The Ghost of Christmas Past in particular helps Scrooge understand his own character (including his ends) by reminding him of the personal history that formed his character.[246] Or we may need to learn "the relative importance of various ends," analogously to learning the relative purchasing power of foreign currencies. The fact that intuition often presents no solution to value

[241] In Ellis's terms, "[v]alues… are basic desires and their objects are basic ends." Ellis, "The Main Argument for Value Incommensurability," 38n1.
[242] Ellis, "The Main Argument for Value Incommensurability," 29–30.
[243] Sen, *Rationality and Freedom*, 609.
[244] Ellis, "The Main Argument for Value Incommensurability," 29–30.
[245] Ellis cites Harris, "Value Vagueness," 158–60.
[246] Dickens, *A Christmas Carol*, chap. 2.

conflicts is compatible with the possibility that a solution might be "worked out"—more on this below.[247]

Another way of moving from value pluralism to value incommensurability is to suggest that "distinct values exclude one another in a crucial sense… each basic value has, *and asserts its importance from*, its own point of view." This exclusion would preclude commensuration along Ellis lines, since commensuration represents an all-things-considered evaluation. But perhaps, as Steven Lukes argues, certain values like friendship demand "total and one-sided commitment to one point of view."[248] Each end makes unqualified evaluations in its own favor, and "[s]ince each basic desire evaluates options solely by how well they achieve *its* end, no such evaluation admits the relevance of other considerations."[249] In this way, the irreducibility of basic ends is likened by some incommensurabilists to the non-fungibility of interpersonal value. Among individuals, if one gains and another loses, the loser must evaluate the situation negatively and the winner must evaluate it positively; they cannot agree. Persons (and objectives, on the analogy) always contest the neglect of their interests and "don't concede the legitimacy" of rankings opposed to their own. Ellis, like MacIntyre, refers to Berlin's claims that agents "are faced with choices between ends equally ultimate, and claims equally absolute, the realization of some of which must inevitably involve the sacrifice of others" and that "human goals are many… and in perpetual rivalry."[250]

Practical Reason Despite Incommensurability?

Importantly, many incommensurabilists admit certain ways of comparing "conatively mixed" options, enabling justified decisions among such options. They do not imagine agents regularly becoming "stuck" between options like Buridan's Ass, or making entirely random decisions. But none of their proposed ways of resolving value conflicts involves evaluating options through a compound covering value. Rather, the resolution always involves a kind of partiality to certain ends. For example, some values might be given a higher status than others in a hierarchy, perhaps being considered sacred:

> When someone considers, for example, the case where her options are saving a friend and making some money, she refuses to even consider the value of the money… she doesn't

[247] Ellis, "The Main Argument for Value Incommensurability," 30–31.
[248] Lukes, "Comparing the Incomparable: Trade-Offs and Sacrifices," 188.
[249] Ellis, "The Main Argument for Value Incommensurability," 31.
[250] Ellis 31-32 cites Berlin, *Four Essays on Liberty*, 168, 171.

> compare her options with respect to a covering value that has friendship and money as contributory values—that's what the sacredness of friendship is supposed to rule out.

While this incommensurabilist claim precludes explicitly trading off a higher value for a lower value, it does not give sacred values lexical priority across all choices.[251] Higher values are controlling, but only in limited circumstances of salience.[252]

Joseph Raz, taking a different line, suggests that while conatively mixed options *can't* be compared with respect to a compound covering value, they *can* be compared through standards operating within particular social practices.[253] Practices can fix "standards of success which enable one to rank goods of the relevant type."[254] Because the standards depend on a practice, they are arbitrary in the sense that take for granted their particular ranking of goods. But since practices can be "socially pernicious," adopting a value ranking from a practice must be justified by other considerations, like "acculturation and the normality/innovation distinction."[255] And these are pragmatic considerations, not value judgments.

As Ellis notes, Raz's view still countenances "a kind of unreasoned partiality, albeit to socially determined mixes of ends rather than to particular ends," and so "that inherent standards don't really commensurate ends." While social practices can explain how conatively mixed options are sometimes compared on the ground, the "comparisons [Raz] countenances aren't the result of reasoned priorities over ends."[256] Rather, they presume commitment to certain social practices, a commitment with only a contestable pragmatic justification. A practice's mix of ends can be undermined, rebelled against, rejected from a standpoint outside the practice.[257] So, "[t]here still isn't any right answer to the question about which inherent standards apply in which circumstances." Ellis thinks that because this partiality precedes comparisons, it also precludes them.[258] I'll have a bit more to say about the Raz-Ellis dialectic below.

The Main Problem with the Main Argument

According to Ellis, the incommensurabilists' move involves a fundamental misunderstanding of values: "Values don't make unlimited, exclusive, or absolute claims, and so

---

[251] Bradley, "Objective Theories of Well-Being," 230.
[252] Ellis, "The Main Argument for Value Incommensurability," 33–34.
[253] Raz, "Incommensurability and Agency," 85.
[254] Raz, "Incommensurability and Agency," 94.
[255] Raz, "Incommensurability and Agency," 98.
[256] Ellis, "The Main Argument for Value Incommensurability," 34–35.
[257] Raz, "Incommensurability and Agency," 99.
[258] Ellis, "The Main Argument for Value Incommensurability," 34–35.

there is no barrier to forming reasoned priorities over ends…" Rather, basic ends are simply "attributes that make situations attractive (good, choice worthy, desirable, etc.)" Outcomes usually have some desirable features and lack others, and our evaluations take in those facts, without making any particular feature decisive, at least not by default or necessarily. Each positive feature is simply a reason to prefer that outcome, other things being equal.[259]

Ellis rejects the analogy between individual interests and basic ends: "[v]alues aren't like agents with competing views of the good because they aren't like agents at all."[260] In fact, the distinctness and irreducibility of basic desires is a reason to *reject* the incommensurabilists' analogy with interpersonal competition: "My desire for friendship, say, doesn't have any access to considerations about equality, security, or material well-being." Values "are just attitudes with respect to their proprietary objects. As such, they are silent about their relations with other values."[261] One way of seeing this point is to note that we can make neutral judgments about which outcome better realizes a particular value, judgments which don't presuppose any commitments to a particular point of view. These claims represent a particular narrow perspective, indicating by phrases like "with respect to…", leaving open the possibility of different evaluations from other perspectives. Accordingly, attending to—or even choosing to realize—one value over another does not "betray" the unchosen value.[262]

While a value's claim on an agent might be irreducible, it is not thus unqualified. It is because values are not exclusive that multiple value considerations can "add up" in favor of certain options. Multiple *ceteris paribus* evaluations can "accumulate" to "establish the *gross* benefits and costs of different options without thereby denying that any calculation of *net* benefits and costs is possible."[263] Ellis elaborates:

> Sometimes I attend to only one consideration, sometimes both, but in the latter case, I don't see the goals *themselves* as incompatible—I want them both. I may recognize that I am not able to advance more than one end in a particular situation but that certainly doesn't imply that my goals make exclusive or competitive claims.[264]

---

[259] Ellis, "The Main Argument for Value Incommensurability," 34.

[260] See also Pildes and Anderson, "Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics," 2177.

[261] Ellis, "The Main Argument for Value Incommensurability," 34.

[262] Ellis, "The Main Argument for Value Incommensurability," 35.

[263] Ellis, "The Main Argument for Value Incommensurability," 35–36.

[264] Ellis, "The Main Argument for Value Incommensurability," 36–37.

Difficult value comparisons result from features of my circumstances or my self, not the values themselves. Values like health and pleasure may in some sense compete for a place in my life, a measure of my time and attention, but that is because my time and attention are finite, not because each value attempts to crowd out or shout down all the rest.

The Nature of Commensuration

So, upon recognizing the claims of various *ceteris paribus* evaluations, an agent may naturally proceed to form an "overall comparison," "ultimate evaluation," or an "all things considered" ranking. The sanctity of friendship is undergirded, not by a refusal to count the costs of acting as a friend, "but rather in the fact that, were I to count the costs, I would perform them anyway."[265] The sort of single-mindedness implied by an incommensurabilist account is neither admirable nor even rational in every case. But appropriately prioritizing values is both.

> Of course, an agent's value rankings might give her reason
>
> to cultivate her mental facilities to automatically bring to mind her desire for friendship and not her desire for monetary gain in some cases. Done properly, this would allow her to avoid calculation costs while still reliably advancing her overall interests. The trick, of course, is determining *which* basic desires should be activated *when*. A system of priorities is required to come up with the right determination.[266]

Ellis thus admits that commensurabilism raises difficult questions, including how we go about commensurating, whether we should commensurate, and more generally the relationship between action and deliberation.[267] Value commensuration would require agents "to pull together a set of *ceteris paribus* evaluations into a single, *everything-considered* assessment," requiring "priorities over values." Until such an assessment is made, values remain uncommensurated, leaving options incomparable for the agent at that time.[268] Where would an ordered set of priorities come from?

Ellis proposes "experience-guided calibration" as the best available account of value prioritization. In this ongoing process, "people react (favorably or unfavorably) to life situations

---

[265] Ellis, "The Main Argument for Value Incommensurability," 37.

[266] Ellis, "The Main Argument for Value Incommensurability," 37.

[267] Ellis, "The Main Argument for Value Incommensurability," 28, 32n For example, Millgram, "Commensurability in Perspective," 218 suggests that value commensuration is often unwise; "not worth the trouble."

[268] Ellis, "The Main Argument for Value Incommensurability," 39n9 thinks that the following vocabulary is more intuitive than its competitors in the literature: "if you can eventually discover a way to make options comparable/commensurable then they weren't incomparable/incommensurable in the first place—they were uncompared/uncommensurated." I agree.

where certain values are at stake; they can generalize from those experiences to establish priorities for their different values." By experiencing varying degrees of satisfaction with their experiences, an agent constructs an ongoing mental "self-portrait" that integrates all their desires, whether or not they are all activated at any particular time.[269] The priority set represented by this self-portrait emerges through three conceptual stages. Initially, it includes merely ordinal rankings with respect to particular covering values: *x* is better than *y* with respect to V. Then after sufficient experience, an agent learns to make cardinal comparisons with respect to particular covering values on some occasions: *x* is better than *y* with respect to V to degree D.[270] This allows value comparisons to be represented on an interval scale, at least roughly and in principle. A cardinal comparison describes not just *whether x* is better than *y*, but *how much better* it is.

The third stage is the crucial one, which incommensurabilism precludes: "In order to have a net evaluation of a course of action, a person must learn (from experience) how an advantage with respect to one end relates to a disadvantage with respect to another end." Coming to recognize these judgments of proportion is a familiar kind of "self-discovery." Ellis approves of Fred Schick's notion that people evaluate states of affairs directly, in a way analogous to how we react to the taste of salt: "We note the utility of propositions by how inclined or adverse we are to them relative to those about which we feel the strongest… We have a sort of immediate grasp of where we stand on these matters."[271] Ellis extends the metaphor by suggesting that we can taste "not only for a salt but also for an appropriate blend of spices. Perhaps [we] 'taste' for the appropriate 'blend' of ends as well, i.e., for the relative importance of advancing those ends." By experiencing various degrees of "fit" between outcomes and our priorities, in the form of varying degrees of felt satisfaction, we can generalize to approximate rates of substitution for "increments of satisfaction." With repeated observation of tradeoffs between the same ends, we can begin to sense when things aren't going well and perhaps adjust course. Adjustments may be more or less sophisticated, from a version of operant conditioning to an agent's development of "a full-blown theory of herself."[272]

---

[269] Millgram, "Commensurability in Perspective," 217–18.
[270] Ellis, "Multiple Objectives," 328.
[271] Schick, *Having Reasons*, 24.
[272] Ellis, "Multiple Objectives," 329–30.

Critics may object that no real calibration accompanies experience; agents simply make up their priorities "out of whole cloth," and perhaps any combination of ends is good enough to issue in action. If prioritizations are unconstrained by experience, no real considerations will favor one prioritization over another. But surely, Ellis suggests, there are *some* constraints on motivational systems: "people can't just value anything." Ellis cites Foot's suggestion that we simply couldn't find the repeated clasping of hands intrinsically valuable.[273] The range of "particularly human wants" is indeed variegated, but not infinitely so. Our animal nature, in particular, may be thought to constrain that range, as Aristotelians like Foot, Nussbaum, and the later MacIntyre have in various ways suggested.[274] And while this is a characteristically Aristotelian insight, it is by no means foreign to utilitarians like Mill, Brandt, and Railton.[275] The theory of human nature underlying the constraints on motivations will be essentially contestable but not necessarily arbitrary. Natural disappointment with results, and perhaps communicative interaction with other humans like oneself, can constitute a feedback mechanism for individual reasoners constructing their prioritizations.

Ellis notes that even if rates of substitution among basic ends are discoverable in the way just described, it may still be that "people's interests just don't fit together," which would mean that "there isn't any complete and humanly sustainable priority scheme." It is an empirically open question whether basic ends can fit together in only one way, many ways, or no ways.[276]

IV: Moderate Commensurabilism (*Whose Justice? Which Rationality?*)

Interestingly, MacIntyre seems to have tempered his anti-commensurabilism since *AV*. He does not explicitly address his own previous arguments in favor of incommensurabilism, but his change of mind seems to be motivated primarily by an increased loyalty to Aquinas. While MacIntyre's 2007 prologue to *AV* maintains his commitment to that book's "central theses,"[277] he admits to having learned a great deal, primarily from Aquinas, since its publication.[278] He takes this transition to have made him *more* faithful to Aristotle, not less, and this is certainly true with respect to the problem of value conflicts—Sophocles no longer plays a significant corrective role in the tradition of the virtues. Despite acknowledging a variety of irreducible

---

[273] Foot, "Moral Beliefs," 89–98.
[274] Nussbaum, "Human Functioning and Social Justice."
[275] Brandt, *A Theory of the Good and the Right*, VI–VII; Railton, "Moral Realism"; Mill, *Utilitarianism*, II.
[276] Ellis, "Multiple Objectives," 331.
[277] MacIntyre, *AV*, vii.
[278] MacIntyre, *AV*, viii–ix.

goods and aims, MacIntyre now maintains with Aristotle that these various ends can be ordered within a "conception of a single, albeit perhaps complex, supreme good." This good is the *arche* of all practical deliberation and also of ethics as a theoretical inquiry.[279]

MacIntyre now notes that "characteristically modern conceptions of goods" do not allow for such a "uniquely rational way of ordering goods within a scheme of life." No names are attached to these "modern conceptions," but MacIntyre seems to primarily have in mind preferentist strand of utilitarianism and welfare economics:

> Those, for example, who identify a good with any object of desire—or perhaps with any object that would be desired, were the person who desires provided with adequate factual information—cannot but recognize the variety, heterogeneity, and incommensurability of such objects.[280]

Because humans enjoy, desire, or value such diverse things, these modern thinkers must countenance "numerous alternative modes of ordering, in the choice between which there are no sufficient good reasons to guide us." This sketch seems to recall the view attributed in *AV* to Weber and Berlin.[281]

At this point MacIntyre's re-asserts a form of *WJWR*'s central claim: an agent choosing among various schemes of values from a standpoint independent of any of them would have no resources for rational deliberation of any kind. The resources for practical reasoning, including the normative standards that guide deliberation, are acquired from some particular place. Inquiry about the nature of this place and its highest good is politics in the Aristotelian sense, comprising the *polis* as Aristotle understood it, and the communities of giving and receiving into which MacIntyre transforms the *polis* in his later work.[282] MacIntyre notes that

> someone who lacked altogether the kind of training and development which Aquinas—and Aristotle—takes to be required would develop the expression of their desires in a piecemeal, uncoordinated way so that they would come to have in adult life desires which appeared essentially heterogeneous, aimed at goods independent of one another and without any overall ordering.[283]

Aquinas's account of *conscientia* and *synderesis* contextualizes his treatment of value conflicts. Agents who "confront the dilemmas constitutive of tragedy" are those whose erring *conscientia*

---

[279] MacIntyre, *WJWR*, 134, 165.
[280] MacIntyre, *WJWR*, 133.
[281] MacIntyre, *WJWR*, 133.
[282] Most centrally, MacIntyre, *DRA*.
[283] MacIntyre, *DRA*, 179.

has directed them toward an action which is in fact "expressly forbidden by divine law."[284] Though such a person cannot recognize it, they are in a dilemma. And it is crucial that in Aquinas' view, "such dilemmas will always rest upon an underlying mistake." At this point MacIntyre contrasts Aquinas' position with Martha Nussbaum's,[285] which is similar in some respects, though not all, to his own position in *AV*.

MacIntyre notes Nussbaum's attention to "a type of situation portrayed in some Greek tragedy in which, so she suggests, full rationality is compatible with recognizing the genuine binding force, in Aquinas' terms, the *pe se* binding force, of conflicting moral claims." MacIntyre suggests that Aquinas and Nussbaum differ in two key ways. Firstly, Nussbaum rejects "the assimilation of practical conflict to disagreement." She denies that the practical assertions of tragic figures have truth-values and so protects them from contradiction. But Nussbaum also claims that Aristotle can accommodate this central insight of tragedy as she understands it. MacIntyre disagrees, highlighting the problem this would generate for Aristotle's understanding of practical reasoning. Put simply, the premises of practical syllogisms must have truth values, and therefore Nussbaum must have gotten Aristotle wrong.[286] The second key difference between Aquinas and Nussbaum is that while they agree that an agent may *seem* to themselves to face conflicting moral claims, Aquinas, unlike Nussbaum, is committed to imputing error to such a person—the conflict cannot be in God's will or "the nature of things." Apparent moral dilemmas are a result of sinfulness, in one way or another.[287]

Whereas MacIntyre had previously understood human goods as competing in perpetual rivalry and occasionally clashing in tragic conflict, he now talks more gently of "the different types of achievement in different types of situation" and "the different character of the goods which are at stake in different types of situations."[288] The differences between goods, and accordingly between virtues, are no cause for the tragedians' lament. Right adjudication of diverse goods is not impossible; rather, it is exactly what the virtue of justice requires:

> it will be impossible to judge justly, and consequently impossible to act justly, unless one can also judge correctly in respect of the whole range of virtues. The courage and skill required in military actions, the temperateness required in respect of pleasures, the

---

[284] MacIntyre, *WJWR*, 186.
[285] MacIntyre cites Nussbaum, "Practical Syllogisms and Practical Science"; Nussbaum, *The Fragility of Goodness*, 36ff.
[286] MacIntyre, *WJWR*, 187.
[287] MacIntyre, *WJWR*, 187.
[288] MacIntyre, *WJWR*, 106.

> liberality or munificence which deserves well of the *polis* because of the provision of resources for public use, the intellectual and aesthetic excellence which makes this dramatist or lyric poet deserve the prize rather than that, all these may at different times have to be judged rightly if a just judgment is to be made.[289]

MacIntyre still resists the extreme form of the unity of the virtues that he attributed to Plato—he admits that progress in the several virtues may not be evenly distributed in the course of an individual life. Nevertheless, the virtues do not make conflicting demands, and as the passage above implies, someone who has any virtue *in its fullest expression* must also have the others.

Apparent Incommensurability as Distinctly Modern

In MacIntyre's view, the apparent incoherence of human desires is not a condition of human being as such; rather, it emerges from a malformed culture. MacIntyre thinks it is a characteristically modern phenomenon that individual preferences involve "a heterogenous variety of needs, desires, and goods conceived in one way or another." And because the notion of an "overriding good" has been rejected by modern liberalism (as MacIntyre understands it),

> [t]he recognition of a range of goods is accompanied by a recognition of a range of compartmentalized spheres within each of which some good is pursued: political, economic, familial, artistic, athletic, scientific… The liberal norm is characteristically, therefore, one according to which different kinds of evaluation, each independent of the other, are exercised in these different types of social environment. The heterogeneity is such that no overall ordering of goods is possible. And to be educated into the culture of a liberal social order is, therefore, characteristically to become the kind of person to whom it appears normal that a variety of goods should be pursued, each appropriate to its own sphere, with no overall good supplying any overall unity to life.[290]

This particular view may not be apt of all "liberal" thinkers or agents, but as MacIntyre notes, they are certainly apt of Rawls, the textbook liberal, who claims that the human self has essentially heterogenous aims, and to "subordinate all our aims to one end" is to "disfigure" the self.[291]

This is a typically MacIntyrean thesis: sweeping, bold, and disturbingly plausible (and accompanied by fewer citations than might be hoped for). Value incommensurability as a practical and philosophical problem is not the same for every human across time and place. The difficulty, and apparent impossibility of value comparison are partially contingent phenomena, the ubiquity and urgency of which increase when, where, and to the extent that a culture fails to

---

[289] MacIntyre, *WJWR*, 106.
[290] MacIntyre, *WJWR*, 336–37.
[291] Rawls, *A Theory of Justice*, 554.

educate its members into some conception of an overall good.[292] Kahneman's empirical work may be considered as an explanation of the agents produced by modern cultures, and to the extent he judges them irrational, a normative critique of those agents.

V: Analysis

It should be clear, now, that the earlier MacIntyre offers a version of what Ellis will call the "main argument" against value commensurability. While *AV* doesn't explicitly analogize conflicts between values to conflicts between individuals, it does picture each value's claim on an agent as absolute, one-sided, and unqualified. Accordingly, MacIntyre's account of tragic conflict is overstated—perhaps too dramatic. Consider again the following passage:

> *both* of the alternative courses of action which confront the individual have to be recognized as leading to some authentic and substantial good. By choosing one I do nothing to diminish or derogate from the claim upon me of the other; and therefore, whatever I do, I shall have left undone what I ought to have done.

A commensurabilist along Ellis' lines will have no qualms with the passage until its very last sentence, which relies on the faulty exclusivist understanding of goods. Obviously, authentic goods may "exclude" each other in a thin sense; I may be in a situation where achieving both is impossible, so I will *miss out* on the neglected good. Certain goods sometimes cannot occupy the same space at the same time in my life. And this fact may be deeply disappointing or troubling; it deserves the word "tragic." But why insist that the neglected good ever made an unconditional claim on my attention? Why not rather see each good as making a *ceteris paribus* claim—if my situation were appropriately different, I certainly ought to pursue the neglected good. But the neglected good doesn't "say" anything about the other value.

MacIntyre's earlier way of moving through value conflicts is similar to Raz's proposal that agents can sometimes be guided in deliberation by value rankings adopted from particular social practices. After all, it is first within the context of social practices (in MacIntyre's semi-technical sense) that *AV* first situates an account of normative virtue and practical reasoning. Ellis's concern with Raz's view was that it permits a kind of "unreasoned partiality" to the contingent value-standpoint of the practice. This critique raises a central issue for MacIntyre's

[292] Admitting this historical-explanatory claim doesn't commit one to agreeing with MacIntyre that the modern liberal condition is lamentable or even problematic. Perhaps the dangers of a socially established conception of the good outweigh the dangers of a lack of such a conception—a possible conclusion emerging from the conflicts of the Big Shift. Here, then, is one angle from which to approach the central question at hand between MacIntyre and his most serious critics—namely, the explanation and evaluation of the transition into modernity.

thought. In *AV* he initially hoped to constrain the moral contingency of practices by situating their function in the narrative context of an individual life and tradition. The three concepts of practice, narrative unity, and tradition were intended to replace Aristotle's "metaphysical biology" with a new teleology to order the virtues. So in *AV*, MacIntyre adds to Raz's picture the claim that engagement in practices is not justified merely as such, but rather, as constrained by the integrity of an individual life which integrates the various roles through which one engages in practices.

As we saw, the earlier MacIntyre suggested that a tragic agent's choice could be better or worse, independently of the chosen alternative, by doing "both what is better for him or her *qua* individual and *qua* parent or child or *qua* citizen or member of a profession, or perhaps qua some or all of these."[293] Such claims have intelligibility with respect to the agent's the context of X's life as unified whole. But making an evaluation that simultaneously respects "some or all" of my various roles, just is a comparative evaluation of outcomes or alternatives. As Ellis says,

> my assessments of outcomes as a citizen, for example, might be incompatible with my assessments as a parent, for example. It doesn't follow that citizenship and parenthood or incommensurable, however, even for me. I can still commensurate those values by making their treatment of shared contributory values consistent."[294]

Incommensurabilism would undermine the picture of decision-making according to which agents act to best realize some set of goals according to some set of beliefs: "There is no *best* achieving a set of goals if, in the absence of priorities, there is no way to compare certain options with respect to those goals."[295] The idea that people act to best achieve their aims is not only a widely intuitive framework for interpreting and evaluating behavior; it is also central to both Aristotelian and utilitarian practical philosophies.

For utilitarians, the association of rationality with choosing "the best" with respect to given aims underlies both the construal of morality as the maximization of universal happiness and the construal of rationality as the maximization of own's own happiness. If options are sometimes incomparable with respect to one's happiness (understood as a comprehensive covering value for all one's aims), then the claim about rationality fails both as a descriptive and

---

[293] MacIntyre, *AV*, 224.
[294] Ellis, "The Main Argument for Value Incommensurability," 41n21.
[295] Chang 1997, Ellis 2006 Ellis, "The Main Argument for Value Incommensurability," 27–28.

as a normative claim.[296] Ellis's claim is merely that if certain goods are *incommensurable*, then *the* good cannot be maximized.

Classical utilitarians, and many non-classical utilitarians, accept a claim that is similar in form but much stronger: if certain values are *irreducible*, then *the* good cannot be maximized. And as I suggested in Part II, the motivation for this stronger claim is the utilitarian commitment to a strictly empirical, quantitative conception of goodness. The options faced by deliberators must, for the classical utilitarians, have expected utility scores that combine cardinal probabilities and cardinal utilities, both of which represent empirical measurements. Ellis's view concurs except with the final clause—utilities can be cardinally-ordered, but not by empirical measurement of a single quantity. A value prioritization represents the relative weights of multiple (irreducible) contributory values, and it is essentially contestable and open-ended.

The ability to choose the best option given certain goals is just as crucial for Aristotelians as for utilitarians. In the textual commentary accompanying his translation of *NE* I, Irwin nicely captures why rational decision-making requires *both* the distinction between ends and means *and* the ranking of non-instrumental ends:

> If we have no idea of what we care about for its own sake and what we care about simply for the sake of other things, we cannot make reasonable decisions about what to do. Similarly, if we simply have a list of things that we care about for their own sake but we have no idea of their relative importance, we cannot make reasonable decisions. A conception of a final good is a conception of the relative weight of the different things that we care about for their own sake."[297]

Like Ellis, Irwin allows that individuals may not have explicated their conception of this final good, but they implicitly rely on one when choosing to pursue any good. For his part, MacIntyre has weakened the language of value conflict from *AV*, recognizing the need for an ordered hierarchy of goods that is knowable, at least in principle, by decision-makers. Whereas he had previously joined Sophocles in blaming tragic value conflicts on reality and human nature, he now takes incommensurability to be merely apparent. Like Ellis, he now blames agents' ignorance.

It is worth noting that Ellis and MacIntyre seem to differ as to exactly *what* a maturing agent is coming to know—for Ellis, value prioritization is a kind of self-knowledge, perhaps

---

[296] This claim is not incontestable. See "Introduction," 2, 13–14.
[297] Aristotle, *NE*, xvii.

constrained in various ways by one's essential nature. MacIntyre often speaks in somewhat more realist or externalist terms about the ordered "goods of excellence" to which virtuous humans devote themselves.[298] But as I noted, MacIntyre still joins Foot, Nussbaum, and Ellis in recognizing how human nature may constrain human ends. So Ellis's process of experience-guided calibration seems well-suited to the Aristotelian picture of an inclusive end that accounts for all of an agent's basics goods, weighted in proportion to each other. Such a value prioritization constitutes the unity of a person's moral life, both in Ellis and in MacIntyre.

In his comments about the "liberal social order," across all his work including *AV* and *WJWR*, MacIntyre also sheds light on the social-historical conditions that generate the philosophical problem of difficult value comparisons. What his critical analysis provides is a standpoint from which to view the scholarly dialectic about value incommensurability. As Ellis notes, a person who has not ordered their own values and developed a unified self-conception will regularly confront difficult value comparisons—values may be uncommensurated, rendering any relevant options incomparable. If MacIntyre is correct, modernity is full of such agents and actively produces them, so we should be unsurprised that the incommensurability thesis is so plausible for many philosophers. Ellis argued convincingly that the inference to the incommensurability thesis from merely uncommensurated values is fallacious. But if the very concept of value prioritization (in Ellis's terms) or overall ordering of goods (MacIntyre's terms) is foreign to agents, comparisons will seem impossible in principle, not just in fact.

A Concluding Proposal

I conclude with a more direct summative statement of the account of reasoning through value conflicts that I think has emerged. Given a choice between multiple outcomes, we must first ask to what extent each outcome realizes this or that specific covering value. These questions may already be difficult; they may even seem impossible. But no choice has been rationally justified until outcomes have been compared with respect to all covering values, then these values are weighted and treated as contributory values to some inclusive good. We may identify this comprehensive covering value with various expressions, like "value-as-such," "goodness," "what is good without qualification," or "the good and the best."

[298] But MacIntyre explicitly acknowledges *both* aspects of the knowledge we acquire about out desires and their objects through practical living and learning. MacIntyre, *ECM*, 3.11, esp. 162.

But in any particular choice, the covering value need not be quite so abstract. I need not decide whether martial courage or artistic beauty is unqualifiedly better. Value conflicts are primarily practical, not theoretical problems.[299] As a practical reasoner, I need not conceptualize any "general way in which A and B trade off in the whole range of situations of choice and comparison in which they figure,"[300] or a "true ranking of the realization of one value against the realization of the other value in terms of one common value across all conflicts of value."[301] This would be Bentham's Enlightenment ambition. Rather, practical reason requires only that "in any given conflict of values, there is a true ranking of the realization of one value against the realization of the other value in terms of some value."[302] Practically speaking, I never need to compare values in the abstract; I always need to decide what is best *for me here and now*.[303]

So it is perhaps most appropriate to identify the comprehensive covering value as "choiceworthiness." It is the final and thus tautologous covering value for every choice. Asking which of my options is most choiceworthy is asking what I have most reason to do, or simply what I should do, full stop. We may proceed to describe my practical reasoning in either Aristotelian/MacIntyrean or decision-theoretic/Ellisian vocabulary:

How can I compare outcomes with respect to overall choiceworthiness?	How can I determine what is best for me, here and now, *qua* human?
I treat choiceworthiness as a compound covering values with several relevant contributory values.[304]	I ask what is best here and now *qua* husband, *qua* teacher, *qua* friend, etc.

[299] Stocker, "Commensuration and Public Reason."

[300] Wiggins, "Is Incommensurability Vagueness?," 59.

[301] Nien-hê Hsieh and Andersson, "Incommensurable Values," sec. 1.3 cites; Richardson, *Practical Reasoning About Final Ends*, 105.

[302] Nien-hê Hsieh and Andersson, "Incommensurable Values," sec. 1.3.

[303] Each practical comparison will be indexed to a person, place, and time. Because agents' judgments are indexed, they may diverge without contradiction, even when the agents are in apparently similar situations. And here, I owe at least a sentence to Kant and the Kantian moral tradition. The generalizability of reasons entails that if another agent were in a *relevantly similar* situation, they ought to make the same judgment. See, e.g., Nagel, *The Possibility of Altruism*, chap. 7. But relevant similarity may be highly specified, so it will remain contestable whether a given decision expressed an appropriate ordering of values, or whether one decision "contradicts" another. But all this, I suggest, is consistent with the idea that values can in fact be compared.

[304] Only a few values will be relevant to any particular outcome comparison, but in principle there may be dozens of relevant values, and the limit case would make all my values relevant.

I assign decision weights to each contributory value, in the form of a cardinal prioritization. (These weights have emerged over the course of my life through the process of experience-guided calibration.)	I formulate a rank-ordering of the various goods that have a place in my life. (My rank-ordering of goods emerges as the goal of the narrative quest that is my life, which I hope to clarify over time.)
In light of my overall value prioritization, I can adopt or participate in an artificial structure, like a schedule (time budgets), to relieve myself from constantly deliberating with respect to the widest possible frame. Scheduled activities may eliminate the need for any deliberation, or they may just function like a decision-frame, restricting the range of values that need to be kept salient during deliberation. Naturally, schedules are often socially coordinated for cooperative focus on shared values, as they are in practices.	By participating in a community that coordinates activities and distributes responsibilities, I can occupy certain stages and roles that focus my activities on the pursuit of particular goods. E.g., when I'm young, I spend much time learning; when I'm acting as a teacher, I focus on the goods of teaching. These social norms suggest default behaviors to avoid the necessity of constantly deliberating all the way up to my conception of the good-for-humans, and all the way back down to the-good-for-me.
As I move from one scheduled activity to the next, I need not be understood as a temporally fractured self, sequentially dominated by various incompatible interests.	As I transition from one stage to the next, or from one social role to the next, I need not be understood as merely taking up various masks, or sequentially adopting opposed hierarchies of goods.
I can aspire to be a unified self pursuing an inclusive aim, which represents the most satisfying balance of values, the best available "self-portrait."	My conception of my good can constitute my unity as a self, expressed in my life narrative, which has coherence to the extent that I possess the virtues of constancy and integrity.

Given my inadequate self-knowledge, developing my conception of this aim will be no easy task, and I am likely to continue encountering difficult choices throughout life.	Given especially the conditions of modernity, I am likely to achieve this imperfectly and so will continue facing choices in which it is unclear what is best.

In placing these tenets in parallel, I do not assert that each item maps onto the other perfectly, as if Aristotelians and utilitarians have been saying exactly the same thing about practical reasoning all along. The structural similarities merely demonstrate how these two paradigms and vocabularies may clarify and extend each other in ways that I have only begun to explore. Aristotelians may need to temper their commitment to value plurality with a recognition of commensurability, becoming comfortable with the formality cardinal value rankings. Utilitarians may need to temper their commitment to comparability with a recognition of irreducibility, becoming more comfortable with the contestability of those same value rankings. Can this be achieved without loss of essential "doctrines"?

I am at least not the first to think so. In Paul Kelly's defense of utilitarianism against MacIntyre, he suggests that commensurabilism can accommodate MacIntyre's distinction between internal and external goods in practices. On MacIntyre's construal of utilitarianism, "[t]he good of any practice must be the instrumental external good of how far it contributes to utility or pleasure." Since MacIntyre's practice involve distinctive and non-arbitrary goods, he thinks instrumentalizing and commensurating them through utility "can provide only the weakest caricature" of their value, "eviscerating their social meaning and personal significance."[305] Kelly doubts that commensuration is evisceration:

> utilitarians often attach a considerable value to practices and institutions which provide the conditions of interest formation and satisfaction, and this involves cultivating and sustaining attitudes to these practices which do not regard them as merely instrumental, but as constitutive of individual well-being. This still leaves the goods of practices as instrumental in promoting well-being, but by regarding practices and their internal goods as constitutive of well-being, the utilitarian's attitude to them can be seen to be much closer to MacIntyre's position.

[305] Kelly, "MacIntyre's Critique of Utilitarianism," 139.

Utilitarians need not conceptualize utility as a hypervalue along Bentham's lines; they may "regard well-being or welfare [as] a more complex concept approaching the Aristotelian notion of 'eudaemonia'." And if this is so, then

> commensurability does not involve the problem of looking for one particular value in all practices and activities, and one which is identified as valuable independently of those practices. This leaves a sufficient degree of instrumentalism in appraising practices, but only sufficient to allow us to prioritize practices and their internal goods."

Kelly also notes that prioritizing the goods of practices along these lines "is something that MacIntyre and the Aristotelian wish to do by appeal to the good and the moral law."[306] Despite MacIntyre's rejection of monistic "utility," his Thomism commits him to some value commensurability, in which "we can discriminate amongst these socially constituted goods: not all goods are *Good*."[307] And again, these discriminations might involve, to put it bluntly, some math. Since expected utilities involves probabilities (represented by cardinal numbers), choosing the best might require knowing how much better one outcome is than another.

If Kelly is right that neither commensurability nor irreducibility are fundamental compromises—"heresies"—on either side, then the problem of value conflicts constitutes not an area of necessary divergence between Aristotelianism and utilitarianism broadly, but an area for productive dialogue between them that results in conceptual progress for both.

[306] Kelly, "MacIntyre's Critique of Utilitarianism," 139–40.
[307] Kelly, "MacIntyre's Critique of Utilitarianism," 136.

3. Aristotelianism, Utilitarianism, and the Gap Between Apparent and Actual Value

Introduction

In this Chapter, I develop a particular axis of contrast between Aristotelian and utilitarian accounts of well-being. Utilitarians are generally reluctant to criticize individuals' attitudes about their own good, maintaining a relatively close connection between what *seems* good to a given person (what they want or enjoy) and what *is actually* good for them. Aristotelians, on the other hand, are generally less reluctant to criticize individuals' attitudes, allowing for a substantial gap between a person's *apparent* and *actual* good—hereafter, The Gap.[308]

This tension between critical and uncritical accounts of well-being plays a central role in the dialectic between Aristotelianism and utilitarianism. The story, and my paper, goes like this: in the first movement, I discuss the critical Aristotelian view of human dispositions (Section 1). Specifically, Aristotelians do not consider what a given person wants or enjoys, just as such, to be a reliable guide to their own good. I note that this perfectionist position generates a cluster of related accusations—elitism, parochialism, dogmatism, alienation—which variously put in question the standards by which Aristotelians criticize individuals' attitudes for failing to track their well-being.

In a second movement, utilitarians seek to replace perfectionism with relatively uncritical forms of hedonism and preferentism (Sections 2 and 3, respectively). By identifying well-being with either pleasure or with the satisfaction of desire, the critical Gap and its vices seem vanquished. But these initially straightforward well-being accounts face their own plausibility objections. And I argue that in the attempt to defuse these objections, utilitarians often reintroduce The Gap in a subtler form. Surveying several qualifications, nuances, and sophistications on hedonism and preferentism, I suggest they undermine the theories' apparent parsimony. Neither the pleasant nor the desirable can be straightforwardly equated with well-being. Criticism of taste and desires seems unavoidable; the gap seems inevitable.

---

[308] Two quick clarifiers: (1) My generalizations about the two traditions may be more or less apt of particular thinkers within them. For example, J. S. Mill's conception of well-being may be more critical than Martha Nussbaum's. But the generalizations hold, and as I say below, they hold *because* of how the two traditions conceive of ethical theory as such. (2) The apparent/actual spectrum cuts across the standard ways of sorting accounts of well-being. While hedonism and preferentism seem naturally less critical and objective list and perfectionist theories are naturally more critical, particular accounts within these categories may fall at various points on the axis. There are more and less critical versions of hedonism (e.g., Epicurean vs. Cyrenaic), preferentism (e.g., laundered vs. actual), and even objective list theories.

I propose a third and final movement, in which utilitarians and Aristotelians both aim to avoid the extreme poles on the apparent/actual axis (Section 4). Both highly critical and highly uncritical accounts of well-being are liable to dangers that are objectionable by both traditions' standards. That is, Aristotelians want to avoid parochialism, dogmatism, and alienation. And utilitarians want to avoid allowing utility to become the normatively empty "fiction" that MacIntyre accuses it of being. Utilitarians and Aristotelians can together inhabit the space between extremes even as they continue to dispute the "size" and metaphysical nature of the apparent/actual Gap. Pursuant to my larger project, it matters that the two traditions inhabit this dialectic space *in full consciousness*.

I: Aristotelian Perfectionism: A Complex Good

Central to any Aristotelianism is the identification of well-being with functional excellence. For a living thing to flourish is for it to exercise the capacities distinct to its kind. So, a good human life achieves the human function, which Aristotelians understand to involve the exercise of reason.[309]

Aristotelians understand the exercise of reason, and a life governed by reason, to be supremely worthy of *desire*. All humans desire eudaimonia—a good life characterized by finality, self-sufficiency, and completeness—and the life of reason best satisfies this fundamental aim. Aristotelians also expect humans who accomplish this function to *enjoy* doing so. Pleasant and painful experiences play at least two roles in Aristotelian ethics. First, they are *indicators* of the quality of one's present activities or circumstances. Put plainly, pleasure is what proper functioning feels like.[310] Secondly, Aristotelians may understand pleasure and pain as themselves *elements* or *constituents* of the quality of one's life. But the contribution of pleasure to *eudaimonia* will not simply be as one more good thing *alongside* good activity and independent from it. Irwin renders Aristotle's cryptic comment on this point as follows: "Pleasure completes the activity—not, however, as the state does, by being present <in the activity>, but as a sort of supervenient end, like the bloom on youths."[311] Pleasure does not merely *accompany* or even *follow from* the goodness of good activity, it somehow *completes* it.

---

[309] As I noted in Chapter 1, I interpret Aristotelianism to accommodate various construals of "the exercise of reason."

[310] This is Stephen Ellis's way of putting it (personal conversation).

[311] Aristotle, *NE*, X.4.8, 1174b35-37. See Irwin's commentary on the passage, pg. 350.

So although Aristotelians do not define eudaimonia fundamentally in terms of pleasure or desire-satisfaction as such, it is not therefore an open question whether the good life is desirable or enjoyable for humans. While there seem to be many humans who do not desire, and would not enjoy, the exercise of reason, such humans simply could not live good lives without some (re)formation of character. A human needs proper dispositions to fulfill their natural potential. A human cannot act virtuously, contemplate, or rightly relate to God (as the case may be) until their affections fit them to do so. So, desire and enjoyment can be expected to accompany *eudaimonia*. Pleasure tracks *eudaimonia*, and desire tracks the prospect of *eudaimonia*.

Nevertheless, Aristotelians predict some Gap between a given person's good and what they most desire or enjoy in their present state. Humans aren't propelled to their good by their nature in the way that stones are propelled to the ground by theirs. Virtue is in this way a *second* nature that agents must somehow develop.[312] The role of virtuous character—which involves reason's governance over appetite—is exactly to close The Gap.

Aristotelians can differ regarding the size of The Gap and what it takes to become virtuous. Aristotelians may have varying pessimism or optimism about human nature and the roles of nature and nurture in virtue-development. Aristotle himself understood virtue development to need active social help, leaving little hope that the improperly raised and educated could be *eudaimon*. Aristotelians influenced by Augustine (e.g.) widen The Gap even further, such that no society can by itself adequately dispose humans toward their true end; divine transformation becomes necessary to close The Gap. But the divine may also be sufficient to close The Gap where nothing else could have been for Aristotle; as MacIntyre notes, "the notion of a final redemption of an almost entirely unregenerate life has no place in Aristotle's scheme; the story of the thief on the cross is unintelligible in Aristotelian terms."[313] Finally, there is no reason an Aristotelian could not be more optimistic than both Aristotle and Christian Aristotelians, allowing that perhaps many individuals can develop virtue even without an Aristotelian education.

---

[312] Aristotle, *NE*, II.1.1, 1103a15-35.

[313] MacIntyre, *AV*, 175. Here, MacIntyre uses "Aristotelian" in a way that precludes the Augustinian picture. In context, he seems to just mean Aristotle's own views.

In any case, The Gap generates the primary features of Aristotelianism to which Enlightenment philosophers like Bentham objected. For example, this objection underlies one of Bentham's most famous assertions:

> Custom has in a manner compelled us to make the distinction between the arts and sciences of amusement, and those of curiosity. It is not, however, proper to regard the former as destitute of utility: on the contrary, there is nothing, the utility of which is more incontestable. To what shall the character of utility be ascribed, if not to that which is a source of pleasure? All that can be alleged in diminution of their utility is, that it is limited to the excitement of pleasure: they cannot disperse the clouds of grief or of misfortune. They are useless to those who are not pleased with them: they are useful only to those who take pleasure in them, and only in proportion as they are pleased.
>
> …
>
> The utility of all these arts and sciences,—I speak both of those of amusement and curiosity,—the value which they possess, is exactly in proportion to the pleasure they yield. Every other species of preeminence which may be attempted to be established among them is altogether fanciful. Prejudice apart, the game of push-pin is of equal value with the arts and sciences of music and poetry.[314]

From such a perspective, the Aristotelian distinction between *phronimoi* and *hoi polloi*, the morally educated and uneducated, looks elitist. For Aristotle himself, the moral elite coincides with a sociopolitical elite, but this need not be the case for Aristotelians as such. An Aristotelian could hold the aristocracy to have socially malformed dispositions and valorize the "natural" human along Diderot's lines. Both master and slave moralities, e.g., are 'elitist' in this sense. But in any case, Aristotelians do not give the pleasures and preferences of each human equal ethical significance.

Naturally, this generates friction with a deeply egalitarian morality like utilitarianism. The utilitarian dictum that *each counts for exactly one* primarily regards the aggregation of well-being, but it may also inform utilitarian conceptions of The Gap. Another person's pleasures and desires ought not count as worth more than mine on any arbitrary grounds of class, education, race, gender, etc. I propose that Aristotelians cannot entirely escape the accusation of elitism. But in some sense, this need not disturb them, because no substantive account of well-being can entirely escape the accusation either. As I show, The Gap always sneaks back into the accounts of well-being intended to eliminate it. This means that some dispositions *do* count more than others on the grounds of a sort of education.

---

[314] Bentham, *The Rationale of Reward*, bk. III, Ch. 1.

The concern with elitism may, then, may not be the mere distinction between the morally cultivated and uncultivated, but rather the apparent arbitrariness or inflexibility of standards for making that distinction. Hence, the related accusations of *parochialism* and *dogmatism*. The character-evaluative criteria for belonging to the moral elite are inevitably culturally particular, emerging within a community in some particular time and place. Perhaps Aristotle merely captures the contingent standards of goodness for Athenian gentlemen, rather the universal standards of goodness for humans.

Aristotelians will accept the contingent *origins* of their evaluative standards but not the contingent *legitimacy* of those standards. Aristotle was quite self-consciously offering an account of the good for humans, of any sort in any place. Aquinas is explicitly anti-parochial in asserting the universal accessibility of the principles of the natural law. Nussbaum and MacIntyre, in different ways, aim to capture our species-nature without overdetermining life plans to fit parochial norms.[315]

Dogmatism may be thought to follow from elitist parochialism to the extent that a community's situated critical standards are not recognized *as* situated and instead treated as capturing timeless and complete moral truths. The smaller The Gap, the more sensitive accounts of well-being will be to empirical evidence. When uncritical theorists ask whether things are going well for me, they will be highly attentive to how I feel things are going. If I don't feel things are going well, this will constitute evidence that some *means* ought to be adjusted. By contrast, if The Gap is large, then the moral elite may frame any objection to their account of well-being as mere evidence that the objector's dispositions are malformed, rather than as evidence that their account may be too narrow. MacIntyre offers an Aristotelian framework that explicitly disallows such attempts to escape dialectical confrontation between traditions. In this framework, rational justification within a tradition requires the confrontation of the strongest available challenges from rivals.

II: Utilitarianism Against Aristotelianism

Utilitarians opposed elitism with egalitarianism, treating all humans' pleasures and preferences as *prima facie* morally relevant. They opposed parochialism with universalism, proposing standards intended to appeal to rational beings as such, not rational beings of a

[315] MacIntyre, *DRA*; Nussbaum, "Human Functioning and Social Justice."

particular and situated kind. They opposed dogmatism with free-thinking, allowing for experimentation and innovation in the pursuit of happiness.

In what follows, I trace some particular dialectics from classical and recent utilitarian literature regarding hedonism and preferentism. I canvas a few objections to hedonism and suggest that adequate rebuttals to them seem to undermine pleasure's status as an empirically accessible ultimate good. I then work through various construals of preferentism, showing how they assume particular normative accounts of human nature.

Why these two, given that some utilitarians have endorsed other accounts of well-being? A maximally egalitarian, universal, and open-ended ethical theory required elimination of The Gap, the Aristotelian contrast between human nature (specifically, tastes and desires) as it is and as it ought to be. No perfectionist, functionalist speculations about human nature, no objective lists. So, classical utilitarians could only build the conception of the human good from what humans already found good—that is, what they in fact enjoy or desire. Moreover, the "facts" about what people enjoy or desire would need to be universally rationally compelling, not depending on contestable interpretation or theory-laden observation. They would need to identify something that, as a matter of empirical fact, seems good to everyone.

IIa: Hedonism: A Simple Good

Pleasure seems to fit the bill. It is perhaps worth noting that some philosophers have talked and still talk as if well-being hedonism is a default component of utilitarianism. This is not true of the utilitarian tradition as I interpret it. Utilitarian orthodoxy includes the claim that we morally ought to do what is useful, understood exclusively as the impartial promotion of well-being. The hedonistic account of well-being is merely one way of giving substance to that fundamental claim. It has attracted utilitarians, especially the classicals, because hedonic quantity seemed to provide an uncritical and empirical standard for well-being judgments. But the natural utilitarian attraction to hedonism (or preferentism) is expressions of the utilitarian aspiration to universal objectivity in ethics.

In any case, hedonism about well-being, or prudential hedonism, asserts that only pleasure has intrinsic value, and only pain has intrinsic disvalue for individuals. Net pleasure alone is intrinsically relevant to individuals' well-being. But what exactly are the mental states called "pleasure" and "pain"? On the standard view, hedonic mental states are *feelings*, *sui generis* qualitative experiences. These experiences have *causes*, but not *intentional objects*.

Sensory pleasures and pains are the most obvious such experiences, but hedonists, including most utilitarians, often insist on the importance of "pleasures of the mind."

On an alternative view, pleasure and pain are attitudes (pro- and con-, respectively), which do have intentional objects. We are variously pleased or displeased *with* or *about* various states of affairs. A hedonism built on this understanding of pleasure, like Fred Feldman's, becomes difficult to distinguish from desire or preference accounts of well-being.[316] In any case, since I aim to make parallel arguments with respect to both hedonism and preferentism, it is not important here whether "attitudinal hedonism," for example, counts as "true" hedonism or collapses into preferentism.[317]

Hedonism identifies well-being with something uncontroversially desirable and empirically accessible. Pleasures are experiences, and everyone likes them. Bradley notes that even anti-hedonists cannot plausibly deny that (1) at least some pleasures increase our well-being, and (2) most other things we consider good can be linked to pleasure in one way or another. And pleasure is a feature of human life that is empirically accessible in the obvious sense that pleasure and pain are felt experiences (perhaps specifically sensory experiences, perhaps not). Hedonic experience is immediately accessible in phenomenal consciousness. It may even give rise to indubitable judgments—I perhaps cannot be mistaken about whether I am experiencing pain or pleasure.[318] For an enlightenment Englishman, such experiences looked to be a promising foundation on which to build evaluative claims.

But these advantages do not attach to hedonism as straightforwardly as utilitarians might have hoped. A plausible formulation of prudential hedonism must be qualified in ways that open a given person's tastes to some kind and degree of criticism. Well-being theorists must allow a gap between what *feels* best and what *is* best.

---

[316] Bradley suggests that whereas some philosophers have construed pleasure as an experience that the subject would prefer to prolong, such a condition seems neither necessary nor sufficient for the experience of pleasure. Bradley, "Objective Theories of Well-Being," 220.

[317] Lin, "How to Use the Experience Machine," 332n28 See also Serena Olsaretti, 'The Limits of Hedonism: Feldman on the Value of Attitudinal Pleasure', Philosophical Studies 136 (2007), pp. 409–15, and Michael J. Zimmerman, 'Feldman on the Nature and Value of Pleasure', Philosophical Studies 136 (2007), pp. 425–37; Lin, 'Pluralism about Well-Being', Philosophical Perspectives 28 (2014), pp. 127–54, at 140–5, 146–7.

[318] Though it might be indeterminate. I may be unsure whether I am experiencing pain or pleasure (or both), but if I think I'm experiencing pain, I am.

Problem 1: Valuable Things Besides Pleasure

Perhaps all humans desire pleasure, but not all humans desire *only* pleasure. Plausibly, we desire some other things for their own sake—perhaps achievement, relationship, or activity. Many of us seem to intrinsically value certain states, not just certain experiences. The locus classicus in establishing this point is of course Nozick's experience machine. Even if we could get much more pleasure by sacrificing realness, we wouldn't. Even if we cannot tell the difference "from the inside" between being in the machine and being in reality, we care which situation we're in.[319]

Hedonists may object that those who would refuse to enter the machine have confused care-for-realness with care-for-the-*status-quo*. Empirical surveys according to which most people would refuse to enter the machine are unreliable, because they mask a *status quo* bias.[320] Respondents are asked to give up their own life, to which they are already attached, by entering the machine. In an alternative dilemma, they are instead asked: suppose you discover that your *present* life has in fact been an experience machine simulation all along—would you *re*-enter the machine? Many more respondents are more friendly to the machine in this scenario.

Hedonists may instead insist that those who would refuse the machine are confusing prudential and moral value. They may bite the bullet and accept that life in the experience machine really is unimpeachably good for the experiencer. Rejecting the experience machine amounts to sacrificing one's own well-being, perhaps in favor of relationships or knowledge, understood as moral goods, or in any case non-prudential goods.

Nevertheless, Lin argues persuasively that despite much confusion about the experiment and its implications, one central point goes through. She clarifies what exactly the experiment should ask us to consider.[321] We compare one life containing certain good experiences which are not real, and a life indistinguishable from the inside, except that the experiences are real. The central question is whether the latter contains any intrinsic value that the former lacks. She suggests the affirmative answer. For example, perhaps I'd rather *feel like and be* a great athlete than *merely feel like* one.

---

[319] This Nozickian objection would apply to any mental-state account of well-being, hedonist or otherwise.Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 122.

[320] Kawall, "The Experience Machine and Mental State Theories of Well-Being"; Brigard, "If You Like It, Does It Matter If It's Real?"; Weijers, "Nozick's Experience Machine Is Dead, Long Live the Experience Machine!"

[321] Lin, "How to Use the Experience Machine."

This is certainly the central point, but it may be worth noting that if realness is a potentially tie-breaking value, it has *some* weight comparative weight against other values.[322] But how much? To answer this question, we must vary the thought experiment by giving the compared world different quantities of pleasure to see how "reality" trades off with enjoyment. Suppose I ask myself whether I'd rather *feel like and be* a **good** athlete than *merely feel like* a **great** one? It matters whether realness is simply a tiebreaker with no comparative value, or whether it has nontrivial value that could compensate for a practically significant quantity of enjoyment. It matters, that is, as a component of the practical project of commensuration that value pluralism makes necessary.

In any case, what matters for my argument is not to conclusively establish the intrinsic value of some activity other than pleasure nor to come to any particular conclusions about how enjoyment and realness trade off. I mean merely to undermine the classical utilitarian claim that pleasure constitutes a good in terms of which human well-being can be uncontroversially and exclusively defined. The dialectic generated by Nozick's few paragraphs has at least established that reflective humans claim to intrinsically value things other than pleasure, such as real activity, relationship, or achievement.

Hedonists are welcome to argue that these rejectors of the machine are confused and in error about what constitutes their good, or even what they themselves value, but it would take serious argumentation to show this. The argumentation must include an account of why things are not what they seem with human desire and well-being. They may tell a story about how certain mistaken value concepts get socialized into machine-rejectors. Of course, the account and story will not be as universally secure as Bentham had hoped. In fact, it will be structurally similar to the story Aristotelians tell about those who fail to desire and enjoy the activities and states that constitute their true good. Both Aristotle and Feldman (e.g.) have accounts of the human good, and when confronted with the fact that many individuals *think* their good is something other than what the account entails, they offer a psycho-sociological explanation for those humans' mistaken value conceptions.

---

[322] Values with only tie-breaking weight are generally implausible. For example, some philosophers have suggested that fairness is relevant to outcome evaluation, but only as a tiebreaker between two outcomes of equal utility. But if fairness is relevant to outcome evaluation at all, then surely it has some nontrivial weight against some amount of utility.

Problem 2: Non-Valuable Pleasures

Some philosophers object to hedonism that *immoral* pleasures are bad (or at least not good) for us. Feldman uses Judah Rosenthal from *Crimes and Misdemeanors* to draw out the intuition that a person can "have it all," but be morally monstrous such that we hesitate to call his life good. Hedonists like Feldman suspect that another confusion of moral and prudential value is at work here. Perhaps Judah Rosenthal's "pleasures [e.g.] *do* benefit him—and that in fact this partly explains why the situation is so appalling, since he so obviously does not deserve to be well-off."[323]

But those unwilling to bite this bullet may draw on Mill's qualitative distinctions between pleasures, suggesting that immoral pleasures are simply of a less dignified kind. To those who ask the obvious question—how do we know which tastes are "higher," Mill suggests appeal to "the competent judge," that is, the experienced expert. But what constitutes higher pleasures as higher, for Mill, is not the exemplars themselves, but the actualized human nature that they exemplify. The wise have learned through experience that, e.g., poetry is a better fit for human beings than pushpin. Mill's account of well-being thus allows for a gap, even a significant gap, between what presently seems, to a given person, to be good for them, and what is actually good for them. And this opens Mill's account of human nature to charges of elitism and parochialism, especially from my Benthamite undergraduate students. Who is this 19th century upper-class English snob to rank poetry higher than pushpin?

Compare Mill's view to the proposal that Fred Feldman calls "desert-adjusted hedonism." Feldman offers this view, noncommittally, to accommodate the intuition that we can take pleasure in things that intrinsically do not deserve it, or deserve it less. The pleasure taken by a sadist, for example, can be "discounted" according to the "unworthiness" of malicious pleasure. Or the pleasure taken by an occupant of Nozick's experience-machine, can be discounted according to the unworthiness of "false pleasures."[324] Welfare only tracks happiness through appropriate objects of enjoyment, perhaps good things in the right contexts:

> Suppose some person, S, takes intrinsic occurrent attitudinal pleasure in some object, O. Suppose the intensity of that pleasure is 10. Then the raw amount of attitudinal pleasure that S takes in O is 10. But we may in addition take account of the extent to which O deserves to be an object of S's pleasure. If O deserves to degree 1 to be an object of S's pleasure, then we can say that the desert adjusted amount of pleasure that S takes in O is

[323] Bradley, "Objective Theories of Well-Being," 224.
[324] Feldman, *What Is This Thing Called Happiness?*, chap. 10.

> 10 × 1, or +10; whereas if O deserves to degree .5 to be an object of S's pleasure, then the desert adjusted amount of pleasure that S takes in O is 10 × .5, or +5. In general, the desert adjusted amount of pleasure in an episode of intrinsic occurrent attitudinal pleasure is equal to the raw amount of that pleasure multiplied by the degree to which the object of that pleasure deserves to be an object of the subject's pleasure.[325]

(Feldman thinks that pain or displeasure cannot be mirrored with pleasure along these lines, since displeasure reduces welfare regardless of its objects. Undeserved displeasures receive no value discount.) On this view, welfare tracks "desert-adjusted happiness." For his own part, Feldman wonders if the view amounts to a confusion between moral and welfare evaluations; "the term 'the good life' often serves to blur or blend these two forms of evaluation."[326]

I think Feldman hesitates to endorse the view because he doesn't share Mill's normatively thick conception of human nature. Without what MacIntyre calls "a conception of human nature as it could be if it realized its *telos*," claims about desert seem arbitrary. For Mill and for MacIntyre, the "two forms of evaluation" are not so different; like Aristotle's Athenians, they suppose the good life to consist in "living well and doing well."[327] Feldman is more comfortable leaving no gap between the best and the most pleasant, because such a gap seems too contestable.

Non-hedonic evaluations of pleasures significantly complicate hedonism, such that it is no longer true by definition that what I most enjoy, as I presently am, is what is best for me. To achieve my best, I may need to cultivate better tastes—that is, tastes for the higher pleasures or more worthy objects. And of course, this move re-introduces The Gap. What *feels* best is not necessarily what *is* best.

Problem 3: Empirical Inaccessibility

The first-person nature of hedonic experience also results in a problem specific to utilitarian hedonists. The utilitarian agent adopts the final end of *impartial* practical reason as the final end of their *individual* practical reason. But whereas each agent has direct (empirical) epistemic access to their own well-being, they have only indirect access to the well-being of everyone else. The impartial hedonist must rely on secondhand evidence about what others experience.

---

[325] Feldman, *What Is This Thing Called Happiness?*, chap. 10.
[326] Feldman, *What Is This Thing Called Happiness?*, 215.
[327] Aristotle, *NE*, bk. I.4.2, 1095a17.

There are two forms of such evidence: people's *reports* of their experiences, and *observation* of people's behavior. Both sources of information involve considerable complications for empirical accessibility. Experience reports may be unreliable even if "honest," involving no intentional deception. Daniel Kahneman's work helps explain why: even sincere reports of human experience are mediated by the reflective mind. There can be a disconnect of troubling severity between the phenomenal mind that has the experience, and the reflective mind that introspects and reports on it.[328] A second problem with hedonic reports regards interpersonal comparisons. Plausibly, you and I could endure the same quantity and quality of pain or pleasure but report the hedonic level of our experiences differently (even with no intent to deceive and no disconnect between the experiencing and reporting selves).[329]

As for behavior revealing what people most enjoy, the central problem is that people do not always seem to behave according to their own hedonic best. A kind of elitism is again lurking for the hedonist. Since I don't always do what pleases me most, even the hedonist admits a gap between my actual behavior and the behavior that would best conduce to my well-being. How can this gap be interpreted or explained? Given that I act to achieve my ends, I act mistakenly when I misunderstand means, my ends are confused, or both. If I am confused about my own ends, hedonists are welcome to call my behavior irrational, but again, they will need to show that my non-hedonic aims are somehow self-deceptive or unnaturally induced.

Now, even if my behavior is indeed always aimed at pleasure, but it often fails to achieve it, I must not be able to easily recognize the means to maximal pleasure. If it turns out that some kinds of pleasure are qualitatively better than others, or even that some sources of pleasure are quantitatively more useful than others, then it's plausible that some people know better than others how to enjoy themselves, or have tastes that enable more enjoyment. These people's desires and practical judgments ought then be given greater weight than others', and hedonists reintroduce a critical attitude towards individuals' practical aims, even if their ultimate ends are left alone.

Hedonism: Upshots

To respond to their critics, hedonists must sophisticate, and so complicate, an initially straightforward view. In chapter 2, I noted how *pluralistic* hedonism complicates the view by

---

[328] Kahneman, *Thinking, Fast and Slow*, chap. 37.
[329] Kahneman, *Thinking, Fast and Slow*, pt. V., esp. Ch. 37. See also MacIntyre, *ECM*, 161–62.

precluding the collapse of all values into a single measurable quantity. I now note how *critical* hedonism (including qualitative distinctions, desert-adjustment, and a notion of expertise) complicates the view by legitimizing the notion of good taste, where that term presupposes a particular and contestable account of human nature. With the introduction of critical qualifications hedonism simultaneously becomes *more adequate to human nature* and *less able to satisfy Enlightenment demands*.

Hedonism failed and fails to meet the full classical utilitarian aspiration to a construct a universally compelling account of well-being from straightforward facts about what humans enjoy. Hedonism has continued to draw defenders of great stature even to today, but most post-classical utilitarians looked elsewhere, generally to preference-based accounts. I turn to these shortly, but it's worth noting that hedonism must not be dismissed as sub-philosophical and swinish naivety. The English deist Anthony Collins once remarked that atheism would seem less plausible were its opponents not so desperately motivated to refute it. Citing Collins, MacIntyre has made a similar remark about relativism, suggesting that "[n]othing is perhaps a surer sign that a doctrine embodies some not-to-be-neglected truth than that in the course of the history of philosophy it should have been refuted again and again."[330] What goes for atheism and hedonism goes also for hedonism. Every age has vehement critics of hedonism precisely *because* it is a perennially tempting idea. Remember Bradley's observation: despite a long history of philosophical opposition to strict hedonism, pleasure's status as an intrinsic good is perhaps less disputed than any other candidate. Whatever else is good, surely pleasure is good.

Those who accept The Gap need not dispute the goodness, even the non-instrumental goodness, of pleasure. For Aristotelians, pleasure is a fallible signal that things are presently going well. When something pleases us, it reveals what we are disposed to find good. This is especially true of pleasures with intentional objects—when I'm pleased *that* something is the case, it seems good to me. To the extent that I am virtuous, my being pleased at something entails that it is good. To the extent that I am not virtuous, there is a gap between what feels good to me and what is good for me.

But as I have noted, Aristotelians can understand pleasure as more than a mere perceptive signal. Plausibly, it's also an intrinsic and constitutive component of well-being. Aristotelians could concede that pleasure is always good in itself, that it is always good for someone, in at

[330] MacIntyre, "Relativism, Power, and Philosophy," 5.

least one respect, to experience pleasure. Aristotelians may even concede that, all other things being equal, more pleasure is better than less. But since pleasure is not the only component of well-being, there is no strict correlation between hedonic value and well-being. I might live a highly pleasant but bad life. Plato and Aristotle both have arguments that this is improbable—virtue is a better bet on average—but perhaps a tyrant could get "lucky." Aristotelians could admit that in these rare cases, someone who fails to fulfill their function is better off enjoying themselves than being miserable. Overall well-being is always accompanied by pleasure, but pleasure is not always accompanied by overall well-being.[331]

IIb: Preferentism: Goodness Formalized

In what follows, I aim to show that desire-based accounts of well-being, like pleasure-based accounts, cannot escape the criticism of immediate desires. But the grounds for this criticism are not as obvious, universal, or incontestable as many economists and preferentists assume them to be.[332] Like hedonist utilitarians, preferentist utilitarians ought to accept the project of criticizing individuals' attitudes despite the essential contestability of such criticism.

Broadly speaking, desire- or preference-based accounts of well-being have several attractive features, especially by utilitarian lights. I noted above that hedonists cannot easily compare hedonic value either within a person or between persons, which makes it difficult to aggregate value quantities and evaluate outcomes. By contrast, economists suppose preferentism to allow for outcome comparisons, since people's actual choice behavior—empirical phenomena—reveal the relative strength of an individual's desires.[333] Preferentism may be motivated by naturalism, into which objectivist theories (including hedonism) are hard to fit.[334]

---

[331] Alternatively, Aristotelians might adopt Feldman's suggestion to "discount" the value of pleasure according to the "worthiness" of the objects of enjoyment. A life would then *not* rank any better in the final evaluation by being pleasant, because what the non-virtuous enjoy is not worthy, and thus their pleasure has little or no value.

[332] Heathwood suggests that while Hobbes and Spinoza may have endorsed desire-satisfaction accounts of well-being, Sidgwick was one of first to discuss them thoroughly. Preferentism spread during the 20th century, especially among economists and decision theorists, whose models often equate welfare with preference satisfaction and preferences to revealed preferences. Heathwood, "Subjective Theories of Well-Being."

[333] This is true by definition if 'preference' is read as 'revealed preference' rather than 'mental preference'. I will discuss this distinction more shortly, but it is worth noting here that if preferences are instead interpreted as mental entities which may or may not be embodied in action, the empirical accessibility of preference strength becomes more complicated. The revealed preference view, for its part, faces the same challenges as psychological hedonism—it requires a contestable error theory to explain why our actual (revealed) preferences do not match our stated (mental) preferences. See Crisp, "Well-Being," sec. 4.2.

[334] Hedonism is *objective* in that it makes actual states of experience, not experiencers' attitudes towards those states, determine the experiencers' well-being. Desire theories are not objectivist in the same way. Unlike

Hausman and MacPherson note that aversion to paternalism is another motivation to prefer formal to substantive theories of well-being more generally.[335] Finally, preference and desire theories take their fundamental cue from the ineliminable intuition that well-being should have some connection to what agents want or are attracted to.[336] For at least these reasons, desire-based accounts of well-being are presently the most popular among utilitarians, appealing to their aspiration for an objectively normative ethics.

Felt Desires and True Ends

The simplest kind of desire theory would claim that the satisfaction of any of my immediate desires is good for me.[337] If I want something, then it would benefit me to get it, all things being equal. An immediate challenge faces this view. Many of my felt desires are merely *instrumental* (I want to drive, but only because I want to visit a friend), or merely *specificatory* (I want *that drink*, but only because I want *a cold drink*).[338] The satisfaction of instrumental and specificatory desires would seem not to benefit me in any way independent from the satisfaction of the relevant intrinsic desire. When I get *that drink*, I'm not extra well-off just for wanting that drink *and* a cold drink; only the underlying desire seems to matter. Perhaps we can avoid this awkward double-counting of satisfactions by saying I am well off to the extent that my *intrinsic* desires are satisfied.[339]

The distinction between intrinsic and non-intrinsic seems straightforward in the simple cases above. But it is not so in other cases. Is my felt desire to spend time with X, who is a friend of mine, a merely specificatory desire, where the underlying desire is to spend time with 'a friend of mine'? Not exactly, but the opposite is not exactly true either, because if I had never become friends with X, perhaps I would now feel a desire to spend time with some different friend, Y. As

---

other objectivist theories, though, hedonism includes only one item on its "list," reducing the value of things like knowledge, achievement, or friendship. Heathwood, "Subjective Theories of Well-Being," 211–12; Bradley, "Objective Theories of Well-Being," 220, 232; Crisp, "Well-Being," sec. 4.3.

[335] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 120.

[336] Heathwood, "Subjective Theories of Well-Being," 212.

[337] Dissertation readers likely need no reminder that "satisfaction" should be understood here in its logical sense, not its common English sense. The satisfaction of a desire or preference is not the agent's feeling of contentment upon getting what they wanted, it is the fact that they got what wanted. Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 105.

[338] Heathwood, "Subjective Theories of Well-Being," 212.

[339] Mark Murphy is one of few theorists who dispute this point. He sees no need to restrict well-being to the satisfaction of "intrinsic" desires, because he simply doesn't believe in specificatory and instrumental desires. The desires that philosophers identify as such are really just desires that motivate in combination with specificatory or instrumental *beliefs*. Murphy, "The Simple Desire-Fulfillment Theory."

for instrumental desires, I may ask: is my desire to teach students philosophy merely instrumental, where the underlying desire is for students to learn philosophy? Or do I value "teaching students philosophy" at least partially in its own right as an activity? And of course, desires are often embedded in multiple instrumental layers; e.g., I want to drive to Norman because I want to get to my classroom, because I want to teach, because I want students to learn philosophy. The complex nesting nature of ends evidences that intrinsic ends must often be remembered, clarified, or even discovered through reflection.

I am not suggesting questions like these have no answers, only that the answers are not often obvious, and they may not be accessible through simple introspection. It may require a sophisticated form of self-knowledge to grasp how my final ends relate to my specificatory and instrumental sub-ends. There might be a gap (Gap?) between my felt desires and "what I *really* want." Felt desires do not wear their status as intrinsic or not on their sleeve; reflection may be necessary to perceive or even construct that status.[340]

From Desire to Preference

Another problem with immediate desires—conflicts among them—highlights the advantages of the concept of 'preference' over 'desire' in formulating a well-being account. Rational choice theory use the concept to precisify the folk concept of desire and clarify its explanatory functions.[341] Basic folk psychology explains intentional behavior by supposing that my actions result from a *desire* for some end combined with a belief that some action will realize that end. My behavior is intelligible, or 'rational' in a minimal sense, if it aims to satisfy my desire.

If I have two desires (whose ends are irreducible), the folk-psychological picture is barely complexified—I will choose an action that realizes the stronger of my two desires.[342] We may introduce preferences here, since preferences are *comparative* in a way that desires are not—I

[340] Volf et al., *Life Worth Living*, 8–10.

[341] Ellis, "Rational Choice."

[342] Some other mechanism must account for cases where desires appear to be equally strong. Besides flipping a coin or letting someone else decide, I may simply have a capacity to decide randomly (i.e. for no conscious reason). In such cases, there may still be a causal explanation for my choice; perhaps I willingly subject myself to framing effects that give one desire more salience or felt power than the other.

desire an apple, whereas I prefer an apple to an orange. But in such a case, 'prefer' seems to do no conceptual work beyond 'desire more'.[343]

The concept of 'preference' becomes unavoidable if we suppose I can have three basic desires, where the strongest desire favors action A and the two weaker desires favor action B. Perhaps I will simply act according to my strongest single desire, as I did in the two-desire case. But plausibly, the weaker desires may somehow combine to "outweigh" the stronger.[344] The language of desire is inadequate here, because to say that I act to get "what I most want" underdetermines, or underexplains, my behavior. We can say, instead, that I do what I *prefer*—I aim for the outcome that best realizes my ends, taken together. In any given choice scenario, my preferences can be understood as a ranking, allowing for ties, of alternative states of affairs.[345] Since such a preference ranking captures my various desires within a specific moment, we can call the view that my well-being consists in the satisfaction of my immediate preferences *synchronic* preferentism.

This model of agency plausibly captures how some agents *in fact* make decisions under multiple objectives. But crucially, the status of this model of agency as *normative* is not incontestable. Ellis calls this an "Aristotelian" model which conceptualizes my agency as somehow unifying the plural ends that compose my psychology. By contrast, a "Heraclitean" model of agency predicts that my choices *always* simply express the dominance of my most salient end in a given moment.[346]

On one hand, these two models present alternative explanatory models, alternative answers to the question, "how does an agent choose given multiple ends?" Plausibly, the answer to the explanatory question varies according to the agent and even particular choices. Some choices are "Heraclitean" and others "Aristotelian." On the other hand, if preferences are to play the central role in an account of well-being, then a model must be chosen as normative for

---

[343] Not all scholars treat the concept and phenomenon of 'desire' as somehow more basic than that of 'preference' in the way that I do here and in what follows. It may be worth noting that Von Neumann and Morgenstern are among those who don't. See Schroeder, "Desire," sec. 3.3; Pollock, *Thinking about Acting: Logical Foundations for Rational Decision Making*; von Neumann and Morgenstern, *Theory of Games and Economic Behavior*.

[344] Along with my own discussion of value plurality, see Ellis, "Multiple Objectives."

[345] Reiss, *Philosophy of Economics*, 37 This ranking is only hypothetical, since of course no agent consciously forms a ranking of every possible outcome of their choice options. But importantly, the decision-theoretic axiom of completeness requires that the agent could in principle form a preference between any two outcomes.

[346] Ellis, "Multiple Objectives."

prudent action. And this choice will be contestable, since it is neither a "matter of fact" nor a "relation of ideas" that to act in an "Aristotelian" manner leaves a person better off. This is so, in turn, because what it means for a person to fare well naturally depends on what a person is. If human agents are bundles of psychological drives, along Hobbes' or Hume's lines, then why prefer a democracy of desires over the tyranny of a single strong desire? But if human agents are substances to which drives belong, then perhaps their flourishing requires the unification of initially disparate and conflicting drives.

Economists and rational choice theorists assume the "Aristotelian" model to the extent that they measure welfare through preference satisfaction *per se*. I've retained scare-quotes for Ellis's use of the term to preclude the sense of a hasty jump to Aristotelian conclusions on my part. I am indeed trying to show what economists and utilitarians must concede to Aristotelians in formulating a defensible account of well-being, but this is merely the first step in that direction. What I want to conclude at this point is minimal, though not trivial. To equate my well-being with the satisfaction of my preferences, rather than the satisfaction of my strongest immediate desire, requires thinking of rational agency (or *me*), at the moment of choice, as potentially something above the expression of any particular desire.

Consider Crisp's case of the "angry adolescent," which can be leveraged against both synchronic preferentism and diachronic preferentism.

> This boy's mother tells him he cannot attend a certain nightclub, so the boy holds a gun to his own head, wanting to pull the trigger and retaliate against his mother. Recall that the scope of theories of well-being should be the whole of a life. It is implausible that the boy will make his life go as well as possible by pulling the trigger. We might perhaps interpret the simple desire theory as a theory of well-being-at-at-a-particular-time. But even then it seems unsatisfactory. From whatever perspective, the boy would be better off if he put the gun down.

Crisp suggests that the conflict is not merely between present and future desires, but plausibly between multiple present desires, some 'foregrounded' and some 'backgrounded'. Most utilitarians seem to agree with Crisp, endorsing something like the "Aristotelian" model.

But the normative claim and the interpretive model do not follow from the formalized theory of rational choice. They are not self-evident. If "Aristotelian" choice among plural ends is rationally normative, it is because human agents are the sort of agent which, in order to flourish, must stand above any particular desires within their preference set. The boy's desire to retaliate against his mother belongs to *the same being* as his other (backgrounded) desires—for survival

and the achievement of other ends. This claim is so widely taken for granted as to seem uninteresting. But it is neither incontestable nor trivial. It matters deeply that the boy's desires all belong to one thing—an animal, perhaps—which flourishing only through integrity as an animal despite radical desire conflicts.

Synchronic Preferentism

So, first-pass preferentism about well-being can be formulated as the claim that my good consists in the satisfaction of my preferences, where my preference set unifies as a diverse desire collection. This is still a relatively bare view, clearly embodying preferentism's central appeals: empirical accessibility (since behavior is at least evidence about preferences), universal attractiveness (since preferences capture what I'm attracted to), and neutrality with respect to the good (since only the *consistency*, not the *content* of my preferences is evaluated).

On a revealed preference account of well-being, accusations of imprudent action can only be justified by reference to faulty beliefs. If I act mistakenly, I must have either miscalculated probabilities or else been misinformed about the relevant facts. (Imprudent action should be distinguished from bad luck—If I make a winning bet but still lose, my action left me worse-off, but I didn't act imprudently.)

By contrast, synchronic preferentism enables a new sort of criticism. The standpoint of critique is that of my whole preference set—the critic is *me*, conceived as a synchronically unified agent. Synchronic preferentism is more plausible than the strongest-desire theory exactly because its conception of human nature goes beyond any self-evident axioms or obvious facts. That conception makes psychological integrity within a moment normative. The authority of any given desire has been challenged; a gap, even if a subtle one, has appeared.

Diachronic Preferentism

Synchronic preferentism is a desire theory adjusted in response to the incoherence of my desires in a given moment. A second coherence challenge faces the preferentist: my desires change over time, and thus my preferences often change over time.[347] If I have incompatible preference sets at t1 and t2, which is in my best interest to satisfy? Heathwood notes that most of us are comfortable discounting or even ignoring past desires in favor of present ones, even when the past desires were "long held and strong".[348] R. M. Hare proposed counting only "now-for-

[347] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 124.
[348] Heathwood, "Subjective Theories of Well-Being," 216.

now" and "then-for-then" desires as relevant to well-being. But this criterion might exclude too much. If at t2 I deeply regret a choice which satisfied my t1 preferences, it would plausibly have been better for me to choose differently despite my preferences.[349] Moreover, we sometimes act in order to satisfy the preferences of the dead, at the very least imagining ourselves as acting for their sake.[350]

If "my" preference set across time is not identical to my preference set at any particular moment in my life, then there is a gap between what seems good at *t* and what is good at *t* for me. The salient desires that make up my preference ranking are subject to criticism from a perspective that transcends each temporal moment. What is that perspective, and whence its authority over *me-now*? Since preferentists insist on value-neutrality, they will note that the only source of desire critique here is ultimately *me*. The only thing wrong with my immediate preference set is its incompatibility with my future preference set. By what criteria might future preferences outweigh immediate ones? The same criteria by which one immediate desire might outweigh another—the relative strength of desires.

But again, it's not obvious that an individual's considered preferences, or lifetime-summed preferences, should take priority over their immediate ones. The commonplace belief that they should presupposes a contestable normative conception of transtemporal unity. That is, they presuppose a certain concept of what a human agent is. Such a conception insists, for example, that the angry adolescent is *the same being* as the one whose life may be cut short by an impulsive decision. Again, it matters deeply that the boy's desires, present and future, all belong to one animal, which persists as an animal throughout significant developmental change and radical preference shifts. Again, to the extent that such a conception *is* normative, it will be contestable.

The basic thought that agents ought not practically contradict themselves is worked out at length in the axioms of decision theory. These are at least worth stating here, briefly and roughly:
Transitivity:

If I prefer A to B and B to C, I prefer A to C.

Completeness:

Given A and B, I either prefer A to B, prefer B to A, or am indifferent.

---

[349] Heathwood, "Subjective Theories of Well-Being," 216.
[350] Heathwood, "Subjective Theories of Well-Being," 216.

Continuity:

> If I prefer A to B and B to C, then there is some unique probability P such that I am indifferent between a lottery between
>
> A with probability P and C with probability 1 – p
>
> and
>
> B with certainty.
>
> In other words, small changes in outcomes do not generate "jumps" in preference rankings.

Independence of Irrelevant Alternatives:

> If I prefer A to B when the two are presented together, then I don't prefer B to A when presented a third alternative.[351]

In my view, these axioms accurately capture minimal criteria for rational choice given a preference set. They are not logically contestable, but their applicability to human agents is contestable. Some thinkers point out that ordinary people often violate the axioms without seeming obviously irrational—e.g., Allais defends the subjects of his paradox as rational.[352] Tversky called such defenders "lawyers for the misguided."[353] But Allais' sympathetic line of thought makes sense if we simply plug a different "unit" of analysis into the theoretical framework. If I violate the transitivity axiom through a preference cycle, there is no logical reason why I should not simply be interpreted as multiple agents, temporally indexed, who simply disagree about the best choice. Each agent is rational.

Other philosophers join Kahneman in defending the axioms' normativity. If my actions violate the axioms, I act irrationally. But to show that I have violated the axioms, another must describe the set of reasons I should've acted from, and those reasons will come from the perspective of my whole preference set.[354] The analysis will be that (1) from the perspective of my whole preference set, I acted irrationally, (2) from the perspective of some subset of my

---

[351] This is my own heavily de-formalized construal of widely available material. See, of course, Steele and Stefánsson, "Decision Theory."

[352] For a representative discussion, see Allais, "Le Comportement de l'homme Rationnel Devant Le Risque: Critique Des Postulats et Axiomes de l'école Américaine."

[353] As reported in Kahneman, *Thinking, Fast and Slow*, 314. For a representative discussion, see Slovic and Tversky, "Who Accepts Savage's Axiom?"

[354] Davidson, "How Is Weakness of Will Possible?"; Ellis, "The Varieties of Instrumental Rationality."

whole preference set, I acted rationally, (3) I ought to have acted on my whole preference set instead of a subset. The first two claims are a matter of interpretating and applying the choice axioms. The third claim gives normative weight to integrity, presupposing the unity of an agent across time. The sense of 'ought' it uses is the rational or prudential sense, asserting that I would be better off having acted otherwise. This is only so, if *I* am something other than a bundle of my desires at many moments.

So the conception of well-being as the satisfaction of axiom-respecting preferences is not so thin and incontestable as it might first seem. My immediate desires, even all at once, cannot be taken as a strict guide to my good, and they can be critiqued from standpoint external to each particular set. Perhaps this standpoint is external merely to my present moment—my past, future, or whole-life. But crucially, the agent supposed by the standpoint doesn't really exist in the way that the agent-of-a-moment does. Rather, it is a normative ideal, not an empirical notion, though I may actually approximate the ideal in moments of "clear-headedness." But crucially, my ideal is constructed as I learn how to interpret and evaluate my behavior and its consequences. And since that learning process is bound to be dialectical and social, others' interpretations of my own behavior may be built into the "objective" atemporal standpoint of desire-evaluation.

Diachronic preferentism is more plausible than merely synchronic preferentism, and much more plausible than immediate-preferentism and the strongest-desire theory. But it is more plausible exactly because it goes beyond any self-evident axioms or obvious facts, subjecting human agency to a normative conception of integrity across time. It's plausible to most people because most people accept the conception of human nature it presupposes. But that conception further challenges the authority of any given desire; it widens the Gap.

Idealized Preferentism

But the satisfaction of even synchronically and diachronically consistent desires doesn't easily equate to well-being. For a desire within such an axiom-respecting set may still be ignorant, malformed, frivolous, or malicious. And plausibly, it would not be good for individuals to satisfy such desires, even if a coherent preference set could include them. Some preferentists therefore "discriminate among preferences,"[355] excluding certain kinds of problematic desires from their accounts of well-being. Some define well-being in terms of a hypothetical condition

[355] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 126.

that somehow idealizes my desires through full information, proper formation, etc.[356] In each case, the exclusion or idealization widens the gap between actual immediate desires and the individual's good.

Preferentists may avoid attempt to avoid appeal to idealization by insisting that such cases all ultimately reduce to incoherence problems. Perhaps what is wrong with ignorant, malformed, frivolous, or malicious desires is that they conflict with other, presumably stronger and more numerous desires. E.g., if a desire compels me to a poisonous drink or futile life-project, perhaps I can't be healthy or safe. If my preferences were made consistent, I wouldn't pursue that end anymore, and so satisfaction of my coherent preference set is still good for me.[357] While this reply may be apt of many cases, there seems to be some badness left over even after making a preference set coherent. Besides, as I have argued, coherent-izing a preference set already involves a normative ideal that transcends any empirical realities, including actual desires and actual agents.

Idealized Knowledge

The most moderate restriction comes in the form of a knowledge condition.[358] If I have a desire based on a false belief, it may not benefit me to satisfy my preferences; e.g., if I long for a glass of mineral spirits that I think is water.[359] Or if I lack certain true beliefs, my preference set may not include a desire that it would benefit me to satisfy. If I don't know that free tacos are being handed out just across the street, I won't want to go across the street. Thus, some philosophers redefine well-being as the satisfaction of my preferences under conditions of idealized knowledge.[360] It would be good for me to cross the street, because if I knew about the delicious distribution I *would* want to go.

There is, of course, a great deal of literature on this topic, on which I simply aim to make two comments. By way of introducing the first: Mark Murphy objects to idealized-knowledge conditions that it is misleading to understand this idealization as a correction of my desires *per se*. In neither given case is something wrong with my desires as such, the problem is entirely with my *beliefs*. The desires that seem to require idealization, like my desire for a glass of

---

[356] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 68.
[357] Heathwood, "Subjective Theories of Well-Being," 214.
[358] Rational or well-informed preferentism: Gauthier 1986, ch. 2; Goodin 1986; Griffin 1986.
[359] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 124.
[360] Heathwood, "Subjective Theories of Well-Being," 212–13.

mineral spirits, do not in fact exist. What other theorists call "basic desires" are, in fact, our only desires, and the "instrumental desires" and "specificatory desires" from which "basic desires" are commonly distinguished do not, in fact, exist. They are simply desires combined with beliefs which attach them to particular instances or instruments. For Murphy, I may be *motivated*, by a combination of a desire and a faulty belief, to do something contrary to my own best interest, but I do not thus *desire* anything contrary to my own best interest.

But Murphy's separation of desires from beliefs looks too strict; our desires often respond, either immediately or gradually, to what we take to be true. They generally depend on what we think is possible or probable, and they may even depend on our beliefs about what other people want, or what's *worth wanting*. When desires are not responsive to reality as we understand it, we often consider psychiatric help.[361] So the preferentist cannot confine rational criticism to beliefs in contrast to desires. It is no more justifiable to scrutinize beliefs instead of desires if some scrutinized beliefs may *be* or at least *imply* evaluations.

Informed preference accounts of well-being raise another question: what exactly is full information or idealized knowledge? What determines the range of facts to which desires are accountable?[362] The answer is surely not *all facts*, because not all facts are practically relevant. But even the range of practically relevant facts is enormous: facts about choices, options, consequences, facts about self and others, facts about possible worlds and counterfactuals. Arguably, idealized knowledge may result in radical changes in values that alienate the idealized agent from the actual agent.[363] Consider the difference between what it would be rational for me to desire given that some version of Islam or some version of Naturalism more adequately described reality. The knowledge condition raises highly contestable questions about the grounds for preference critique. Neither the axioms of rational choice nor any universal observations about human desire can plausibly determine the range of relevant information.

Idealized Desire-Formation

Some philosophers define well-being as the satisfaction of preferences that are *not malformed*. How is appropriate desire formation to be understood? One way of drawing the contrast between well-formed and malformed desires is by reference to 'natural' and 'unnatural'

---

[361] Brandt, *A Theory of the Good and the Right*, VI.

[362] Heathwood, "Subjective Theories of Well-Being," 214.

[363] Rosati, "Persons, Perspectives, and Full-Information Accounts of the Good"; Sobel, "Full-Information Accounts of Well-Being."

desires. MacIntyre recounts that in Diderot's famous dialogue, Rameau challenges the *philosophe* with "the question of precisely which of our desires are to be acknowledged as legitimate guides to action, and which on the other hand are to be inhibited, frustrated or re-educated."[364] As MacIntyre comment on the passage, he notes that desires themselves cannot be the criterion, because "all of us have, actually or potentially, numerous desires, many of them conflicting and mutually incompatible." To decide between rival motivations, we must somehow "order a variety of impulses, felt needs, emotions and purposes." So practical reason's criteria, including moral rules, cannot "be derived from or justified by reference to the desires among which they have to arbitrate."[365]

Once a preference set is specified, the general principle that "more is better" with respect to desire-satisfaction can do some action-guiding work. All things being equal, it is better to satisfy *stronger* desires and to satisfy *more* rather than fewer desires. But of course, when things are not equal—when the quantity and intensity of desires favor different choices—a nontrivial normative commitment to synchronic integrity must adjudicate. And then, if rival preference sets can be shown to be equally coherent, a nontrivial normative commitment to diachronic integrity must adjudicate.

As MacIntyre notes, Diderot's own answer to the question distinguished "natural" human desires, exemplified by his mythologized Polynesians, from "artificial" desires implanted in us by society. But that very distinction precludes Diderot's "attempt to find a basis for morality in human physiological nature." If the desires we in fact experience and observe cannot be trusted as such, then the critical criterion for desires cannot come from human nature as it actually is.

MacIntyre argues that Hume faces the same problem, despite his "philosophically more sophisticated account" of social morality. Hume agrees with Diderot that moral judgments arise from the passions rather than reason. But because "in judging morally we invoke general rules," moral rules must somehow be useful in satisfying the passions. But whose passions? Hume's view, like Diderot's, involves "an implicit, unacknowledged view of the state of the passions in a normal and what we might call, but for Hume's view of reason, reasonable man." For MacIntyre's Hume, "[t]he normal passions are those of a complacent heir of the revolution of 1688," and Hume disdains groups like the Levellers, Jansenists, cynics, and ascetics as deviant

[364] Diderot, "Rameau's Nephew," 14ff and passim.
[365] MacIntyre, *AV*, 48.

"enthusiasts."[366] Thus, Hume covertly criticizes the passions with "a highly conservative normative standard," no more justified than Diderot's.[367]

Another strain of thinking contrasts 'autonomy' and 'heteronomy' in desire formation. Many thinkers rightly notice that individuals are often limited by social norms in ways that prevent them from achieving and enjoying all that they otherwise might. As Hausman and MacPherson say, "[p]eople may have particular preferences because of previous coercion or manipulation, and they may come to prefer things as the result of problematic psychological mechanisms."[368]

The autonomy-heteronomy contrast may simply be a specification of the natural-unnatural contrast, since autonomously-formed desires seem prototypically natural and heteronomously-formed desires prototypically artificial. Perhaps individuals' preferences are not really *their* preferences, or at least not relevant to *their* well-being, unless those preferences are somehow constructed by a robustly free self. MacIntyre's reasoning thus applies here, too. The relevant kind of autonomy cannot mean absolute independence from social relations; it must mean independence from social relations that are problematic in a specified way. In *DRA*, MacIntyre himself sketches the dangers to independent rational agency constituted by parental overreach, insecurity, inflexibility, and the like.[369] The human reasoner raised in MacIntyre's ideal community of giving-and-receiving is indeed granted a kind of autonomy in preference formation, being directed by teachers towards goods that are independent of any particular teacher.[370]

A third and particularly fascinating vocabulary in which the contrast between well-formed and badly-formed desires is the vocabulary of therapy. Richard Brandt, among other

---

[366] MacIntyre, *ECM*, sec. 2.2, esp. 81–84.

[367] MacIntyre, *AV*, 48–49. MacIntyre goes on to point out Hume's failure to justify morality to the sensible knave, who breaks moral rules when he can get away with it. In the *Treatise*, he refuses to resort to "any innate spring of altruism or sympathy for others" in responding to the knave. But he does so resort in the *Enquiry*—"an invention intended to bridge the gap between any set of reasons which could support unconditional adherence to general and unconditional rules and any set of reasons for action or judgment which could derive from our particular, fluctuating, circumstance-governed desires, emotions and interests." Adam Smith resorts to sympathy to bridge the same gap, which is in fact "logically unbridgable," which reveals 'sympathy' as yet another "philosophical fiction" alongside 'rights' and 'utility.'

[368] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 127; Nussbaum, *Women and Human Development: The Capabilities Approach*, chap. 2, esp. 135–148; Superson, "Deformed Desires and Informed Desire Tests"; MacIntyre, *ECM*, 198–99.

[369] MacIntyre, *DRA*, 89–90, 102-106,.

[370] MacIntyre, *DRA*, 70–71.

thinkers, suggested in 1979 that the preferences relevant to my well-being are those which would survive cognitive psychotherapy.[371] The suggestion is particularly interesting in light of MacIntyre's identification in 1981 of the Therapist as one of the "Characters" of modern moral culture. These characters embody, in various social contexts, that culture's loss of "the distinction between rational and nonrational discourse" and "the distinction between manipulative and nonmanipulative social relations."[372] In contrast to the Manager, the therapist's domain is personal, rather than professional life. He "treats ends as given, as outside his scope; his concern also is with technique, with effectiveness in transforming neurotic symptoms into directed energy, maladjusted individuals into well-adjusted ones." Their role is conceived as entirely amoral; they are "uncontested figures, who purport to restrict themselves to the realms in which rational agreement is possible-that is, of course from their point of view to the realm of fact, the realm of means, the realm of measurable effectiveness."[373]

But the authority of effectiveness, for MacIntyre, is vacuous without presumed ends. Of course, every therapist presumes a conception of what maladjusted and well-adjusted individuals believe, desire, and do. One Aristotelian therapist has been quite clear about this—the philosopher and practiced psychoanalyst Jonathan Lear. For Lear, the central philosophical significance of the psychotherapeutic tradition, and the specific practice of free association, is to cultivate an agent's self-understanding, a crucial precondition for self-control—that is, right ordering of the parts of the soul.[374] This tradition and its practices, which MacIntyre himself engaged with at length,[375] are not value-neutral. More generally, any robust therapeutic methodology presupposes some account, perhaps broad and flexible, of what is good for the patient.

Once a well-being theorist is willing to subject immediate desires to the sort of thick critique as that involved in therapeutic transformation, it is hardly a step further to say that the desires that it is good for me to satisfy are desires that conform to virtue. Once desires have become thickly rational—fully informed by available facts, re-formed through therapy in light of

---

[371] Brandt, *A Theory of the Good and the Right*, pt. 1.
[372] MacIntyre, *AV*, 30.
[373] MacIntyre, *AV*, 30.
[374] See, e.g., Lear, "Alasdair MacIntyre and Therapeutic Method"; Lear, *Wisdom Won from Illness: Essays in Philosophy and Psychoanalysis*.
[375] See esp. the 2004 preface to MacIntyre's early book on Freud's theory of the unconscious: MacIntyre, *The Unconscious*.

right reasoning—the perspective from which my desires are critiqued cannot simply be *me* in any straightforward empirical sense.

An Aristotelian might suggest that the standpoint has become that of the virtuous person. But it would be more proper to identify the standpoint as that which I would adopt if I were virtuous. My preferences do not give virtue its normative authority. But what I ought to do is nevertheless conceptually bound to the preferences that *I myself* would have, given the proper knowledge and character formation. Even from my own present position, I can recognize the limitations of my present position. But of course, I come to know my idealized preferences through interaction with other people—parents, teachers, friends, masters in practices, exemplars.[376]

Two Sorts of Utilitarian

What this survey constitutes is not an argument against preferentism, but rather an argument that preferentism must either be vacuous or else combine with contestable assumptions about human nature. Some utilitarians are committed to denying this case, but others might accept it. As I work toward my broader conclusions, it will be fruitful to briefly draw attention to representatives from these two camps.

For example, Chris Heathwood offers a caution against using desire-idealization to rule out "putatively defective desire," like those of Rawls' grass-counter, Moore's destructive drunk, or a child abuser:

> This move runs the risk of turning the desire theory into an objective theory of well-being in subjectivist clothing. It is not open to idealized-desire theorists to claim that part of what it is to be idealized is to desire *the right things*, that is, the things it is good to get no matter your desires. That is closet objectivism. The conditions of idealization must be stated in value neutral terms, and without reference to things that were identified by a belief in their objective welfare value."[377]

Along similar lines, Hausman and MacPherson note many thinkers' resistance to "a [preference] ranking in terms of some theory of what makes human life good, because economists disavow any substantive theory in favor of the view that well-being is the satisfaction of preferences."[378]

Richard Arneson is perhaps the clearest guardian against value judgments smuggled into preferentism. He defends a theory of nonmoral value, or well-being, "according to which the

---

[376] MacIntyre, *DRA*, 70–71.
[377] Heathwood, "Subjective Theories of Well-Being," 213.
[378] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 69.

good for a person is the fulfillment of his (corrected) tastes and values."[379] In Arneson's view, "What determines the value of satisfying a preference is the attitude the agent would adopt toward it after ideal deliberation." Or again, "[i]t is my enlightened judgment of my perhaps unenlightened preferences that determines their value.[380]

As indicated by the "ideal deliberation" and "enlightened" qualifiers, Arneson acknowledges The Gap, and even names it as such:

> I believe that the judgmental component of preference, which implies that error is possible, has to do with the gap between actual preferences, which may be ill-considered, and the preferences one would have after more careful consideration.[381]

But Arneson's Gap is supposed to be minimal. He is not greatly troubled by preference changes, noting that they are often shallow, being based on false beliefs. In many preference changes, "[m]y underlying first-best preferences stay the same."[382] Accordingly, he maintains the right formal procedure will allow the aggregation of preferences over a lifetime.[383] Arneson's idealization also addresses malformed preferences: "the hypothetical ideal deliberation test resolves our doubts" about preferences formed badly, "because the test has us imagine what the person would want if he could think clearly about his preferences."[384] (This clear thinking about one's desires is presumably what a good therapist helps a person achieve.)

Arneson decisively rejects "the perfectionist claim that some preferences can be judged inherently better or worse than others just by inspection of their content and comparison to a perfectionist preference rating scale."[385] Instead, Arneson thinks "[t]he value of satisfying one or another of an individual's preferences depends on its relative importance as judged by the individual herself."[386] Arneson thus approves of Dworkin's political liberalism, or government neutrality with respect to "the question of the good life."[387] His aversion to paternalism shapes, or is perhaps shaped by, his cautious understanding of The Gap:

> In some cases the gap between a person's actual preferences and his hypothetical rational preferences may be strikingly apparent to the sympathetic observer, but in many cases observers will be in a poor position to determine whether any such gap exists, and in that

---

[379] Arneson, "Distributive Subjectivism," 159.
[380] Arneson, "Distributive Subjectivism," 163n, 164.
[381] Arneson, "Distributive Subjectivism," 162–63.
[382] Arneson, "Distributive Subjectivism," 165.
[383] Arneson, "Distributive Subjectivism," 166.
[384] Arneson, "Distributive Subjectivism," 169.
[385] Arneson, "Distributive Subjectivism," 170.
[386] Arneson, "Distributive Subjectivism," 181.
[387] Arneson, "Distributive Subjectivism," 161.

> broad range of cases a presumption of the rationality of an individual's actual preferences should hold.[388]

Arneson means to defuse the sense that The Gap compromises the prima facie authority of the individual, who should be assumed to know their own best in a "broad range of cases." In the political realm, preferences must not be judged "by citizens generally," because this would imply that others' judgments are "a (perhaps rough) indicator of the objective worth of those preferences." Society's treatment of certain preferences as "expensive" cannot be justified without perfectionist judgment.[389]

The question my foregoing argument has already posed to Arneson is: is there any robust sense in which my preferences can be taken treated as exclusively *mine* over against external judgment? A related question returns us to a central point from Chapter 2: why do I take certain attitudes toward certain objects after ideal deliberation? Are there any constraints, from "ideal deliberation" or elsewhere, on the objects toward which we can take positive attitudes, or on the conditions that generates our tendencies to take positive attitudes?

The fact-value distinction operating in Arneson's view is just the kind Aristotelians challenge, and utilitarians (as I've argued) struggle to maintain. Functional concepts are both factual and evaluative, so to be fully aware of them is to accept certain normative judgments. The factual analysis will be much more loaded than subjectivists will want, but complicated facts still are facts. And while Aristotelian ethical inquiry does presuppose certain critical judgments, they *can* be more than arbitrary parochial assumptions. The critique of human desires emerges in community, through ongoing conceptual and empirical analysis of those desires and the goods to which they're directed. Athenians are not permitted to confine this inquiry to a moral elite: "we aristocrats just know what's good for everyone, and that's that." Nor does an elite have authority to impose their vision of the good on others, especially if Mill was right to consider autonomy a constituent of human welfare.

Arneson is representative of many, but not all, utilitarians. As Hausman and MacPherson suggest, although modified- or laundered-preferentism is in principle formal and not substantive, it may

> provide a foot in the door for arguments concerning substantive goods, goods that it is rational to prefer. To take well-being as the satisfaction of informed rather than actual

---

[388] Arneson, "Distributive Subjectivism," 164.
[389] Arneson, "Distributive Subjectivism," 190.

> preferences shifts the emphasis from what people prefer to what is rational to prefer. If it could be shown, for example, that it is rational for everyone to prefer happiness to unhappiness or virtue to vice, then a view of well-being as the satisfaction of rational preferences would lead to a substantive view of well-being as involving happiness or virtue.[390]

Peter Railton is one utilitarian who seems comfortable with that foot in the door. Unlike Arneson, Railton defends "a realist account of non-moral value—the notion of something being desirable for someone, or good for him."[391] This account of well-being involves several conceptual layers. First, I want or desire certain objects or outcomes; they are positively valenced for me. It would therefore be in my *subjective interests* to get what I want. Second, since my wants are frequently flawed, they are insufficient to independently ground desirableness. If I knew better, I wouldn't want myself to want what I in fact want. Therefore, it would be in my *objectified subjective interests* to get what I would want myself to want if I had full factual knowledge of myself and my environment.[392]

Arneson and Brandt stop here, identifying well-being (more or less) with what Railton calls objectified subjective interests. But Railton adds a third notion that he calls my *objective interest*, which is the "reduction basis" of my objectified subjective interests. That reduction basis is simply the set of facts that *make it true that*, if I were fully informed, I would want myself to want such-and-such.

Railton takes this account of well-being to ground ethical inquiry as realist in a particular sense. Moral judgments can be true in a "fundamentally non-epistemic sense," that moral properties are relational but objective, and that any discipline, including ethics, needs to both consult the facts *and* offer normative guidance in light of them: "Moral inquiry is of a piece with empirical inquiry."[393] Railton even suggests optimistically that despite serious and recognizable failures in contemporary moralities, "we know a fair amount about morality." In fact, "Moral properties supervene upon natural properties, and may be reducible to them."

---

[390] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 128.

[391] Railton, "Moral Realism," 10.

[392] *What I would want myself to want* is not the same as *what I would want* if I were fully informed. This is important, as Crisp says, because

> on the informed desire view the subject must actually have the desires in question for well-being to accrue to her. If it were true of me that, were I fully informed I would desire some object which at present I have no desire for, giving me that object now would not benefit me. Any theory which claimed that it would amounts to an objective list theory with a desire-based epistemology. Crisp, "Well-Being," sec. 4.2.

[393] Railton, *Facts, Values, and Norms*, xvi.

Railton denies that his objective realism entails the prescription of single kind of life for everyone or a single set of norms for every society.[394] He also notes that intrinsic value "straddles the objective and the subjective… our opinions don't determine the matter, [but] intrinsic value does always seem to be related to, and realized by, subjects."[395] Railton identifies as a Humean about motivation in the sense that belief alone is an insufficient motivator; desire is necessary for action. But he rejects what he takes to be Hume's view that "if moral facts existed, they would necessarily provide a reason … for moral action to all rational beings, regardless of their particular desires."[396] Reason doesn't compel agents towards "particular beliefs or practices apart from our contingent, and variable, ends."[397] Moral imperatives are applicable to all rational agents, though they only function as a reason for action for some.

In nearly all these respects, Railton's account of well-being approximates the basic features of an Aristotelian account. Railton's "objective interests" correspond to a thick account of human nature that explains *why* a rational human would prefer particular states over others, rather than treating basic ends as arbitrary and given.

In the face of preferentism's challenges, some practically-minded thinkers admit that preference-satisfaction is not beneficial always and as such, but then suggest that an individual's preferences are still "the best *guide* to what is beneficial to her." The response intends to preserve the formality of the theory of well-being—"[r]egardless of what human well-being truly is, the best measure of well-being is the extent to which preferences are satisfied." Of course, laundered-preferences accounts make individual well-being "even less measurable" than it is on actual-preference accounts.[398] Moreover, the deeper problem with the response as a practical is that it is an empirical claim, one which can't be verified or falsified without identifying what well-being is.[399]

But the suggestion need not be merely an unprovable claim about statistical correlation. It may rather be an insistence that well-being theorists at least do justice to the common conviction that the relation of the good to desire and pleasure is not accidental. Railton manages to respect that conviction by making his preferentism "straddle" the objective and subjective.

---

[394] Railton, "Moral Realism," 5.
[395] Railton, *Facts, Values, and Norms*, xiv.
[396] Railton, "Moral Realism," 6–8.
[397] Railton, "Moral Realism," 8.
[398] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 129.
[399] Hausman and McPherson, *Economic Analysis, Moral Philosophy, and Public Policy*, 120–21.

Preferentism: Upshots

I've argued that preferentism cannot plausibly avoid The Gap. Case by case, proponents propose qualifications on the simple desire theory, each of which challenges the authority of felt desires as such, and the grounds for the challenges are contestable claims about human nature and agency. First, the move from felt desire to intrinsic desire involves an assumption about human self-knowledge. Then, the move from desire satisfaction to preference satisfaction, or synchronic preferentism, involves a normative assumption about the synchronic unity of an agent's psychology, privileging "Aristotelian" accounts of the self over "Heraclitean" accounts. Diachronic preferentism involves a normative assumption about an agent's unity across time, privileging the preferences of certain moments over others. Desire-idealizations of all sorts involve contestable normative claims about which desires would survive idealization, regarding which kind of knowledge is relevant to our well-being, what an agent needs to know to fare well. Therapeutic preferentism involves a normative assumption about which desires are consistent with adequate self-understanding.

Whereas preferentists like Arneson are committed to rejecting this thesis for the sake of value neutrality, preferentists like Railton may be open to it. Perhaps the underlying claims about human nature are just the sort of concerns on which Railton would focus empirical investigation. An Aristotelian can agree—again, to the extent that empiricism is taken in an Aristotelian or Millian, rather than a Humean or positivist sense.

III: Conclusion: The Tension of The Gap

After Virtue's central critique of utilitarian well-being accounts is what MacIntyre describes as their "attempt to derive the morality from the psychology," to draw conclusions about what they should do from facts about what they happen to already want. In this sense utilitarianism was a consequentialism without teleology. And as such, it was for MacIntyre not an ethics at all. It is the distinction between "man-as-he-is" and "man-as-he-could-be-if-he-realized-his-essential-nature," that justifies normative claims—specifically, the notion of human potentiality. The gap between actuality and potentiality is the precondition for morality.[400]

Despite MacIntyre's many philosophical transitions since *After Virtue*, his basic frustration with utility remained the same:

---

[400] MacIntyre, *AV*, 52.

> To maximize utility was on an earlier utilitarian view to maximize pleasure and to minimize pain. More recently the thought has been that to maximize utility is to maximize preference satisfaction. Neither formulation pays regard to the fact that what we are each of us pleased or pained by and what we each of us prefer depend in key part upon our prior moral formation, upon how far we are just, courageous and temperate, and therefore disposed to act rightly. How we conceive utility thus depends on our prior formation and commitments, so that it cannot provide a standard independent of them period to propose utility maximization as such as the measure of right action must therefore be a mistake. What, then, are we in fact doing when we make decisions on the basis of cost-benefit analysis as we often do? The answer is that we are always working with some highly determinate and contestable conception of what is to count as a cost and what as a benefit in this or that type of case and with some prior determination of whose costs and who's benefits are to be counted, whose costs and benefits ignored. It is evaluations already made or presupposed that allow us to find application for the notion of utility in our decision making. The notion of utility maximization as a free standing notion that by itself provides guidance for action is a philosophical fiction.[401]

What I take myself to have shown is that MacIntyre's claims are most apt of the least critical forms of utilitarianism, whether hedonist or preferentist. Identifying well-being with pleasure or preference-satisfaction as such and in fact, they leave no gap between the actual (descriptive) and the potential (normative). But when both hedonism or preferentism are qualified in response to objections, The Gap returns. That critical Gap presupposes a standpoint from which actual desires are criticized, and I've argued that this cannot be confined in any sharp way to the agent's own actual evaluations.

Utilitarians hoped to maintain this confinement to avoid both alienation and elitism-parochialism-dogmatism. Against the latter cluster, ethical inquiry was to aim at objectively normative practical conclusions, derived from universally-accessible conceptual truths and empirical facts. Hedonism and preferentism appealed to utilitarians because taste and desire are matters of quasi-empirical fact. Humans enjoy certain outcomes (roughly, we perceive their realization as good) and want certain outcomes (roughly, perceive the prospect of their realization as good). The objects of our desires and the sources of our enjoyment seem good to us. And attention to what we want and like is a powerfully intuitive starting point for inquiry into goodness; just as attention to what *seems* to be the case is the inevitable starting point for inquiry into reality.[402]

---

[401] MacIntyre, *ECM*, 77.

[402] I am inclined to think, but unable to show conclusively, that this parallel can be pressed quite far. DiIulio's J. S. Mill agrees. DiIulio, *Completely Free*.

But as I argued in Chapter 1, ethical inquiry cannot meet the epistemic standards of Enlightenment empiricism. The utilitarian tradition must release the aspiration to a theory of well-being *constructed wholly from empirical facts*. It may preserve some continuity with its origins by aspiring to a theory of well-being *constructed from empirically accountable judgments*.[403] "Value-seemings" will be taken as defeasible *prima facie* evidence about value. Desires and tastes are thus like value-detection faculties, our psychological equipment for perceiving various kinds of goodness. As such, they function like our perceptive faculties—they provide *prima facie* judgments that should be compared intra- and inter-personally to determine what is actually good.

So, any practically objective ethics must recognize a meaningful Gap between what *seems* good and what *is* good. The most robust hedonisms and preferentisms are Gappy in this way, since neither enjoyment nor the satisfaction of desires or preferences can be plausibly identified with the good. And The Gap might be quite large indeed—all of us at least regard the dispositions of childhood in this way, and in principle, an adult could fail to close in all sorts of ways. The version of me that desires my true well-being might be far from the present version of me, even as an adult.

But an alienated conception of well-being is not viable because well-being has a fundamental conceptual connection with desirability. If my good is conceived as something neither desirable nor enjoyable to me, it is *alienated*.[404] A general problem accounts of well-being that ideal my given tastes or desires is that The Gap might be too large for my alleged well-being to interest *me*, as I actually am.[405] The Gap must not alienate well-being from attraction.

---

[403] Some aspects of a minimal core of well-being are empirically measurable. These may include features of biological health or animal well-being, like nutrition, the functioning of organs, and environmental fit. These conditions appear good for human animals, though humans have to learn to perceive this. And even minimal animal well-being will not be measurable until these various components are commensurated. As we saw in Chapter 2, commensurations are contestable, emerging holistically (accounting for judgments about actions, principles, and exemplars) and dialectically (updated continually in response to new experience, information, and reflection).

[404] Since a critical conception of *my* good would presumably be held by *you*, critical conceptions in general may seem to encourage paternalism on your part. But the moral question—whether you know better than I do what's best for me—should be distinguished from the political question—whether you should try to influence me for my own good. A perfectionist may even include autonomy in their conception of well-being, as Mill seems to.

[405] Heathwood, "Subjective Theories of Well-Being," 214. See also Bradley, "Objective Theories of Well-Being," 234.

Theorizing well-being thus involves a tension between crucial desiderata: an account ought to be both motivating and normative. The same dialectical pattern appears in the dialectic among hedonists and among preferentists: the original theory of well-being is selected because it looks maximally thin, specifying the good life formally and without contestable assumptions. But the thinner the account of well-being, the less plausible and action-guiding it can be. Aristotelians are caught in the same tension, but begin from the opposite side. They have thicker conceptions of well-being, which involve robust assumptions, often of an allegedly "metaphysical" character. Those who don't share their parochial assumption charge Aristotelians with alienation and elitism.

There is no way to escape the tension, only better and worse ways of inhabiting it. If an account of well-being proposes a form of life that many apparently rational people do not recognize as good, it will need to be complemented by a careful account of moral formation and perhaps transformation. A substantially critical well-being account will need to suggest some process by which a person who gravely misunderstands their own good, and so doesn't desire it, could come to understand and desire it. This is perhaps a way to "soften the blow" of The Gap—even if I don't now desire or enjoy what is best for me, I may put myself in a position to develop different dispositions. Of course, I will need to be somehow motivated to thus order my life, which is why I might aim, e.g., at virtue because of what it delivers instrumentally, rather than what it is, then learn to value it intrinsically with time.[406]

Critical well-being theorists must also guard against the genuine problem of dogmatic and parochial elitism. Aristotelians and utilitarians alike must not exploit The Gap to appeal to rational standards acceptable merely to some narrowly construed "us," insulating that elite from potential challenges. What might this look like? Consider how Jen Hatmaker, a popular religious author, reflects on the cultural norms under which she was raised:

> I was taught that my body, my feelings, my perception, my desires were fully untrustworthy, that I was absolutely not a reliable source of authority in my own life. In fact, we were taught overtly to outsource all of our authority. So it was up to everybody else to tell us what to do, how to act, what to believe, how to live. In fact, we were taught to actively distrust our instincts. If we really wanted something, that should raise a red flag. That could be perceived as worldly or selfish… particularly for girls.[407]

---

[406] Arguably, Plato adopts or discusses this strategy at various points in his corpus.
[407] Camp, *Jen Hatmaker: When Everything Breaks: Grief, Growth, and Human Flourishing*.

Hatmaker here explicitly describes how belief in The Gap can become a tool for insulating an authoritative moral elite from critique. Hatmaker later narrates how her body often "knew" she wasn't faring well at various points. But she was taught to interpret these sub-conscious value perceptions as evidence that her own dispositions were malformed, rather than as evidence against the authoritative moral norms. Hatmaker's experience was in fact a challenge, even an empirical challenge, to that conception. If an apparently mature and rational person is well-off according to some account of well-being but nevertheless experiences misery and despair, this is important *prima facie* evidence against the theory.[408]

I suggest that MacIntyre's account of tradition-bound rationality can place the crucial guardrails around critical accounts of well-being, primarily by subjecting the very conditions of well-being theorization to criticism within and among traditions. To the extent that any theory insulates itself from criticism, it cannot be affirmed as the best theory available. The open-endedness and invitation of critique built into traditioned inquiry resists parochialism to the greatest extent possible without the Enlightenment attempt to escape all particularity.

[408] Likewise, if an account predicts that X is not well-off, but X consistently feels quite happy this is also *prima facie* evidence against the account, as long as there is no independent reason to doubt X's rationality.

Afterword: What I Have and Haven't Done

I have discussed suggested some points of conflicts between Aristotelians and utilitarians, describing in each case how their respective fundamental commitments may be refined through the conflict. I proposed a conception of utilitarianism as a tradition of moral inquiry in MacIntyre's distinct sense, describing the continuities and discontinuities with utilitarian history that this new conception would involve. Regarding value plurality, I argued that utilitarians and Aristotelians can together acknowledge the irreducibility and commensurability of values, treating commensuration as capacity to be developed for the integrity of practical reasoners. Regarding human dispositions, I argued that criticism is both inevitable and inevitably contestable. Any moral community must work *from* a particular normative conception of human nature, and it ought also to work *on* that conception in response to internal and external challenges to it.

This work opens up many avenues for further exploration. The difficult and unavoidable work of inquiry I laid out in my Preface will have to continue. I've found myself in the middle of a grand tradition; in fact, caught between two grand traditions. I've only started to describe my own relationships to these traditions by attending to parts of their central events, resources, and challenges. The conclusions I've drawn highlight further questions that would make my inquiry more complete.

More Histories

For example, I want to make utilitarianism's history as clear as MacIntyre makes Aristotelianism's history. A narrative that adequately capture utilitarianism would need to say *some* more about Bentham, Mill, and Sidgwick, and *much* more about the utilitarians before and after them. More specifically, I think the pre-classical, classical, and post-classical phases of utilitarianism each deserve chapter-length treatments in MacIntyre-style analysis. Despite the deserved canonical status of the Big Three, they do not exhaust its history. On the later side of the tradition, Kelly notes in his defense of utilitarians against MacIntyre that

> another story could be told featuring the likes of W. S. Jevons, F.Y. Edgeworth, Alfred Marshall, G. E. Moore, R. S. Harrod, and contemporary theorists, such as J. C. Harsanyi, R. M. Hare, Derek Parfit and James Griffin, which takes the utilitarian tradition into contemporary moral and political thought. Such an alternative narrative would illustrate

> how the utilitarian tradition attempts to accommodate the criticisms that MacIntyre raises.[409]

At various points my argument acknowledged this alternative narrative, but it doesn't explicitly narrate it, and someone ought to—from the standpoint of a friendly MacIntyrean. More specifically, the relationship between utilitarianism and the development of metaethics in the 20th century seems underattended. MacIntyre has attended to it by blaming the failure of utilitarianism for the rise of non-cognitivist accounts of moral language. I hope to narrate the relationship differently, blaming non-cognitivism for the failure of utilitarianism. As I read the story, the classical utilitarians were deeply realist about value, but many post-classical utilitarians seemed to lose their realist nerve, preferring an awkward combination of a highly demanding normative theory with a highly subjective account of normativity.[410]

The historical project also demands a fuller reckoning with the significance of Christian theology and practice for the tradition's origins. There is a fairly obvious chronological sense in which Christianity bridges the Aristotelian and utilitarian traditions; it suspiciously occupies the gap between their respective breakdown and ascendency. Attention to Christian ethics, and perhaps its Abrahamic cousins, might clarify my earlier claim that utilitarianism is a splinter tradition from late medieval Aristotelianism. Most broadly, what's necessary is a synoptic history that relates these three specific traditions to ultimately Greek and Hebrew (and perhaps Roman) texts and cultures.

At least one historical continuity entirely distinct from these may provide an intriguing angle on the concerns of this dissertation: classical Chinese philosophy. The parallels between specific texts and thinkers must not be overstated, but I'm convinced the Mohist critique of Confucianism bears fruitful comparison to the utilitarian critique of Aristotelianism and Christianity. Such an observation is commonplace,[411] but I wonder what a MacIntyre-style approach to the Warring States schools could reveal. Interestingly, unlike the classical utilitarians, Mohists did not contrast themselves with Confucians along the lines of religiosity vs. secularity. More generally, I wonder whether the Mohist critique avoids the epistemic mistakes

---

[409] Kelly, "MacIntyre's Critique of Utilitarianism," 131–32.

[410] This awkwardness is perhaps most clearly exemplified in Smart's famous "outline" of utilitarianism, and perhaps most clearly exposed by Williams' "critique." Smart and Williams, *Utilitarianism: For and Against*.

[411] Ahern, "Is Mo Tzu a Utilitarian?"; Vorenkamp, "Another Look at Utilitarianism in Mo-Tzu's Thought."

of which I accuse Enlightenment moralists, and if so, whether the critique lands so as to require fewer concessions to Confucianism in ongoing dialogue.

More Conflict Points

More robust historical work would inform more instances of conceptual rapprochement that I did not address in this work; it is worth mentioning a few of those instances.

I consider it a non-negotiable Aristotelian principle that *activity* be given final, non-instrumental value. I do not think this commitment can admit of degrees or be massaged in any way to accommodate any strict mental-state accounts of well-being, whether hedonistic or ideal. In my two chapters, I aimed not to directly argue against accounts of well-being like hedonism or preferentism. I aimed to show that in a few ways, they cannot be justified or applied on strictly empirical or scientific grounds; grounding and using the accounts both require ineliminably contestable judgments. But if the Aristotelian-utilitarian synthesis will leave behind strict hedonists, then a more serious effort is needed to placate or persuade them. Mill, as usual, provides grounds for centralizing the value of distinctly human *activity* in a utilitarian account of well-being, and I wonder if Moore's conception of "organic wholes" of value may be put to use in conceiving of the interrelated value of desire, pleasure, and action.[412]

I have said less about any particular virtues in these chapters than an Aristotelian might be rightly expected to. MacIntyre suggested that utilitarianism must collapse the *virtues* into a single *virtue*—benevolence, a trait that directly or indirectly maximizes utility. His harsh comments against Benjamin Franklin's account of virtue apply here. But "virtue collapse" is only as necessary as the collapse of irreducible values into a single hypervalue, against which I argued in chapter 2. If human nature admits of multiple irreducible goods, or is susceptible of irreducible kinds of harm, then dispositions to rightly attend to each of these domains of life might constitute distinct traits of character. Of course, to maintain some kind of unity among these virtues (!), governing virtues, perhaps prudence and love, must be introduced. Here, the historical investigation will be of importance in that Christianity introduces the moral centrality of love, which eventually becomes charity in Aristotelianism and benevolence in utilitarianism. My suspicion is that if I am right that no fundamental incompatibility emerges regarding value

[412] Moore, *Principia Ethica*, I.D, 27ff.

plurality and the apparent/actual value gap, then none will emerge regarding the schema of virtues either.

Throughout, I've avoided attaching the word 'happiness' to the final end of practical reason in both traditions, even though it the most common word utilitarians associate with utility, and the most common word used to translate 'eudaimonia'. But the word is so hotly contested as to deserve lengthy comments that I could find no place for in this text. Such Aristotelian heavyweights as Jonathan Lear and MacIntyre disagree about whether it's an appropriate translation of Aristotle's final end.[413] And the place of 'happiness' in the utilitarian tradition is equally unclear—in fact, utilitarians may be partially to blame for the word's ambiguity in English.

I have also said relatively little about one of the central challenges to utilitarianism. For the classical utilitarians, this was the problem of reconciling an egoistic psychology with an apparently highly demanding morality. And despite the release of egoism in later utilitarianism, they still find themselves having to explain a gap between what humans ought to do and what they find themselves motivated to do.[414] I think this is one of many utilitarian challenges that may become more tractable given a more Aristotelian conception of flourishing. In fact, widening The Gap I discussed in Chapter 3, between what is good and what seems good, may *narrow* the motivation gap, between what is good for an individual and what is good from an impartial standpoint. If Aristotelians are right that human nature is constitutively social, then individual flourishing is partially constituted by social flourishing. This is more than the claim that there is some causal relation, even a necessary one, between the flourishing of others and my flourishing. For me to flourish *is* to some significant extent for others to flourish in particular ways. These claims would benefit from elaboration.

Finally, I find myself wishing to clarify the vision of the moral community that my Aristotelian-utilitarian synthesis requires. I have sided somewhat decisively with the Aristotelian view that the study of ethics cannot persuade or motivate those without requisite character and thus cannot benefit them. And I situated this claim within MacIntyre's conception of traditions, which further anticipates the possibility of dialectic impasse between rational persons. But while this amounts to a rejection of a long-standing utilitarian ambition, I take very seriously the

[413] Lear, "Alasdair MacIntyre and Therapeutic Method."
[414] Mulgan, *World, Finitude, Solitude*.

utilitarian sense that ethical inquiry ought to *regard* everyone. Utilitarians aim to eliminate arbitrary moral distinctions, giving appropriate moral significance to all humans as such, all suffering things as such. I don't mean to claim these are *uniquely* utilitarian concerns (MacIntyre has justified a kind of impartial concern for humans as such, and Nussbaum has insisted on attention to animal well-being[415]), but they are *distinctively* utilitarian concerns.

As an Aristotelian-utilitarian, I want *ethics* to be for everyone without expecting *ethical inquiry* to be for everyone. Here again, a more robust history of Hebrew and Greek thought might bear fruit, because the utilitarian attitude strikes me as distinctly Abrahamic. I mean that Judaism, Christianity, and Islam often envision themselves as communities called out from a wider world to advance God's mission in that world. This, despite great variety within and among the traditions regarding the nature of God's mission and the mode of the religious community's participation in it. Why has no grassroots movement of Kantians *as Kantians*, or Aristotelians *as Aristotelians*, achieved the communal status and social significance that Effective Altruism has? It may be a mere historical contingency, but utilitarianism's history as a social reform movement seems as consistent as any other feature of its continuity. Effective Altruism seems to constitute a moral community along the lines of the Abrahamic traditions—accepting the universe's point of view as our own, advancing the practical aims of that point of view in community, and perhaps inviting others to join that mission.

I hope to formulate an account of a moral community, consistent with the commitments developed in these pages, that operates both "in the world, but not of it," and "for the life of the world." I think this is what a utilitarian-Aristotelian moral community would have to be.

These are exciting lines of inquiry that I will be delighted to pursue in future work. My attempt to answer my initial questions has raised many more questions. Being a "lifelong learner" (an unfortunately clichéd expression) means encountering that fact, not with frustration or discouragement, but with gratitude and anticipation.

---

[415] MacIntyre:
I contrasted earlier the particularity of the relationships and responsibilities of family with the universality of the requirement that we respond to the needs of others. But note that this universal requirement is itself always presented to us in particular terms. It's this particular stranger in her or his particular situation whose needs arising from her or his pain or grief I am to address, not those of humanity in general. What's required of us is not universal and egalitarian benevolence, let alone the pursuit of that chimera, 'the great happiness of the greatest number'.
MacIntyre, "Pain, Grief, and Other Signs of Life."

Bibliography

Ahern, Dennis. “Is Mo Tzu a Utilitarian?” *Journal of Chinese Philosophy* 3, no. 2 (1976): 185–93.

Allais, Maurice. “Le Comportement de l’homme Rationnel Devant Le Risque: Critique Des Postulats et Axiomes de l’école Américaine.” *Econometrica* 21, no. 4 (1953): 503–46. https://doi.org/10.2307/1907921.

Anderson, John. *Studies in Empirical Philosophy*. Passmore, 1962.

Andersson, Henrik, and Anders Herlitz, eds. “Introduction.” In *Value Incommensurability: Ethics, Risk, and Decision-Making*. Routledge, 2022.

Andersson, Henrik, and Anders Herlitz, eds. *Value Incommensurability: Ethics, Risk, and Decision-Making*. Routledge, 2022.

Anscombe, Elizabeth. “Modern Moral Philosophy.” *Philosophy* 33, no. 124 (1958): 1–19.

Aristotle. *Nicomachean Ethics*. 3rd ed. Translated by Terence Irwin. Hackett, 2019.

Aristotle. *Politics*. Translated by Terence Irwin and Gail Fine. Hackett, 1995.

Arneson, Richard J. “Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare.” *Philosophy & Public Affairs* 19, no. 2 (1990): 158–94.

Austin, J. L. “A Plea for Excuses: The Presidential Address.” *Proceedings of the Aristotelian Society* 57 (1956): 1–30. https://doi.org/10.1093/aristotelian/57.1.1.

Baier, Annette. “Civilizing Practices.” *Analyse & Kritik*, no. 6 (1984): 61–77.

Baier, Annette. “MacIntyre on Hume.” *Philosophy & Phenomenological Research* 51, no. 1 (1991): 159–63.

Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Batoche Books, 2000.

Bentham, Jeremy. *The Book of Fallacies: From the Unfinished Papers of Jeremy Bentham*. John and H. L. Hunt, 1824.

Bentham, Jeremy. *The Rationale of Reward*. John and H. L. Hunt, 1825. https://archive.org/details/rationaleofrewar00bent.

Berlin, Isaiah. *Four Essays on Liberty*. Oxford UP, 1969.

Blackledge, Paul, and Kelvin Knight, eds. *Virtue and Politics: Alasdair MacIntyre’s Revolutionary Aristotelianism*. University of Notre Dame Press, 2011.

Bradley, Ben. “Objective Theories of Well-Being.” In *The Cambridge Companion to Utilitarianism*. Cambridge UP, 2014.

Brandt, Richard B. *A Theory of the Good and the Right*. Oxford University Press, 1979.

Brandt, Richard B. *Ethical Theory*. Prentice-Hall, 1959.

Brigard, Felipe De. “If You Like It, Does It Matter If It’s Real?” *Philosophical Psychology* 23, no. 1 (2010): 43–57.

Burns, J. H. “Utilitarianism and Reform: Social Theory and Social Change, 1750–1800.” *Utilitas* 1, no. 2 (1989): 211–25. https://doi.org/10.1017/S0953820800000236.

Camp, Lee C. *Jen Hatmaker: When Everything Breaks: Grief, Growth, and Human Flourishing*. No Small Endeavor. Aired September 22, 2025.

Chang, Ruth, ed. *Incommensurability, Incomparability, and Practical Reason*. Harvard UP, 1997.

Chang, Ruth. “Introduction.” In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Harvard UP, 1997.

Conway, Daniel W. “After MacIntyre: Excerpts from a Philosophical Bestiary.” *Soundings* 69, no. 3 (1986): 206–25.

Crimmins, James E. “Jeremy Bentham.” In *The Stanford Encyclopedia of Philosophy*, Fall 2024, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, 2024. https://plato.stanford.edu/archives/fall2024/entries/bentham/.

Crisp, Roger. “Sidgwick and Utilitarianism in the Late Nineteenth Century.” In *The Cambridge Companion to Utilitarianism*. Cambridge UP, 2014.

Crisp, Roger. “Utilitarianism and the Life of Virtue.” *The Philosophical Quarterly (1950-)* 42, no. 167 (1992): 139–60.

Crisp, Roger. “Well-Being.” In *The Stanford Encyclopedia of Philosophy*. 2021.

Davenport, John J., and Anthony Rudd. *Kierkegaard After MacIntyre: Essays on Freedom, Narrative, and Virtue*. Open Court, 2001.

Davidson, Donald. “How Is Weakness of Will Possible?” In *Essays on Actions and Events*, 2nd ed. Clarendon Press, 2001.

Dickens, Charles. *A Christmas Carol*. Chapman & Hall, 1843.

Diderot, Denis. “Rameau’s Nephew.” Project Gutenberg of Australia, 2007. http://gutenberg.net.au/ebooks07/0700101h.html.

DiIulio, John Peter. *Completely Free: The Moral and Political Vision of John Stuart Mill*. Princeton UP, 2022.

Donner, Wendy. “Mill’s Theory of Value.” In *The Blackwell Guide to Mill’s Utilitarianism*. Blackwell, 2006.

Driver, Julia. *Uneasy Virtue*. Cambridge Studies in Philosophy. Cambridge UP, 2001.

Ellis, Stephen. “Multiple Objectives: A Neglected Problem in the Theory of Human Action.” *Synthese* 153, no. 2 (2006): 313–38.

Ellis, Stephen. “Rational Choice.” In *New Dictionary of the History of Ideas*, vol. 5. Cengage Gale, 2004.

Ellis, Stephen. “The Main Argument for Value Incommensurability (and Why It Fails).” *The Southern Journal of Philosophy* 46 (2008): 27–43.

Ellis, Stephen. “The Varieties of Instrumental Rationality.” *The Southern Journal of Philosophy* 46, no. 2 (2008): 199–220.

Feldman, Fred. *What Is This Thing Called Happiness?* Oxford UP, 2010.

Foot, Phillipa. “Moral Beliefs.” In *Virtues and Vices: And Other Essays in Moral Philosophy*, Online. Oxford UP, 2002.

Griffin, James. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Clarendon Press, 1986.

Haidt, Jonathan. *The Righteous Mind: ?* Allen Lane, 2012.

Harris, George. “Value Vaguesness, Zones of Incomparability, and Tragedy.” *American Philosophical Quarterly*, no. 38 (2001): 155–76.

Hauerwas, Stanley H. “Can Aristotle Be a Liberal? Nussbaum on Luck.” *Soundings* 72, no. 4 (1989): 675–91.

Hausman, Daniel M., and Michael S. McPherson. *Economic Analysis, Moral Philosophy, and Public Policy*. 2nd ed. Cambridge UP, 2006.

Heathwood, Chris. “Subjective Theories of Well-Being.” In *The Cambridge Companion to Utilitarianism*. Cambridge UP, 2014.

Hsieh, Nien-he, and Henrik Andersson. “Incommensurable Values.” In *The Stanford Encyclopedia of Philosophy*. 2021.

Hsieh, Nien-hê, and Henrik Andersson. “Incommensurable Values.” In *The Stanford Encyclopedia of Philosophy*, Fall 2024, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, 2024. https://plato.stanford.edu/archives/fall2024/entries/value-incommensurable/.

Hurka, Thomas. *Perfectionism*. Oxford University Press, 1993.

Hyman, Arthur, James J. Walsh, and Thomas Williams, eds. *Philosophy in the Middle Ages: The Christian, Islamic, and Jewish Traditions*. 3rd ed. Hackett Publishing Company, 2010.

Irwin, T. H. "Tradition and Reason in the History of Ethics." *Social Philosophy & Policy* 7, no. 1 (1989): 45–68.

Irwin, Terence. *The Development of Ethics: A Historical and Critical Study*. Oxford University Press, 2007.

Israel, Jonathan I. *Radical Enlightenment: Philosophy and the Making of Modernity 1650–1750*. Oxford University Press, 2001.

Kagan, Shelly. *Normative Ethics*. Westview Press, 1998.

Kahneman, Daniel. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

Katz, Leonard D. "Pleasure." In *The Stanford Encyclopedia of Philosophy*, Winter 2016, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2016. https://plato.stanford.edu/archives/win2016/entries/pleasure/.

Kawall, Jason. "The Experience Machine and Mental State Theories of Well-Being." *Journal of Value Inquiry* 33, no. 3 (1999): 381–87.

Kelly, Paul. "MacIntyre's Critique of Utilitarianism." In *After MacIntyre: Critical Perspectives on the Work of Alasdair MacIntyre*. Notre Dame UP, 1994.

Knight, Kelvin. *Aristotelian Philosophy: Ethics and Politics from Aristotle to MacIntyre*. Polity Press, 2007.

Koons, Robert C. *Is St. Thomas's Aristotelian Philosophy of Nature Obsolete?* St. Augustine's Press, 2022.

Kraut, Richard. "Aristotle's Ethics." In *The Stanford Encyclopedia of Philosophy*, Fall 2022, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, 2022.

Le Mahieu, D. L. "Foreword." In *The Principles of Moral and Political Philosophy*, by William Paley. Liberty Fund, 2002.

Lear, Jonathan. "Alasdair MacIntyre and Therapeutic Method." In *Learning from MacIntyre*. The Lutterworth Press, 2020.

Lear, Jonathan. *Aristotle: The Desire to Understand*. Cambridge UP, 1988.

Lear, Jonathan. *Wisdom Won from Illness: Essays in Philosophy and Psychoanalysis*. Harvard University Press, 2017.

Lin, Eden. “How to Use the Experience Machine.” *Utilitas* 28, no. 3 (2016): 314–32.

Lukes, Steven. “Comparing the Incomparable: Trade-Offs and Sacrifices.” In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Harvard UP, 1997.

Lutz, Christopher Stephen. *Tradition in the Ethics of Alasdair MacIntyre: Relativism, Thomism, and Philosophy*. Lexington Books, 2009.

MacIntyre, Alasdair. *A Short History of Ethics: A History of Moral Philosophy from the Homeric Age to the Twentieth Century*. 2nd ed. University of Notre Dame, 1998.

MacIntyre, Alasdair. *After Virtue: A Study in Moral Theory*. 3rd ed. Duckworth, 2007.

MacIntyre, Alasdair. *Dependent Rational Animals: Why Human Beings Need the Virtues*. Open Court, 1999.

MacIntyre, Alasdair. “Epistemological Crises, Dramatic Narrative and the Philosophy of Science.” *The Monist* 60, no. 4 (1977): 453–72. https://doi.org/10.5840/monist197760427.

MacIntyre, Alasdair. *Ethics in the Conflicts of Modernity: An Essay on Desire, Practical Reasoning and Narrative*. Cambridge UP, 2016.

MacIntyre, Alasdair. “Moral Rationality, Tradition, and Aristotle: A Reply to Onora O’Neill, Raimond Gaita, and Stephen R. L. Clark.” *Inquiry* 26, no. 4 (1983): 447–66.

MacIntyre, Alasdair. “Moral Relativism, Truth and Justification.” In *Moral Truth and Moral Tradition: Essays in Honor of Peter Geach and Elizabeth Anscombe*, edited by Luke Gormally. Four Courts Press, 1994.

MacIntyre, Alasdair. “Notes from the Moral Wilderness [1959].” In *MacIntyre’s Engagement with Marxism: Selected Writings 1953-1974*. Brill, 2008.

MacIntyre, Alasdair. “On Having Survived the Academic Moral Philosophy of the 20th Century.” In *What Happened in and to Moral Philosophy in the Twentieth Century?: Philosophical Essays in Honor of Alasdair MacIntyre*, edited by Fran O’Rourke. University of Notre Dame Press, 2013.

MacIntyre, Alasdair. “Pain, Grief, and Other Signs of Life.” *2nd Annual Fall Conference: “A Culture of Life”* (Notre Dame, IN), Notre Dame Center for Ethics and Culture, 2001. https://www.youtube.com/watch?v=TOTpu9pPn1Y.

MacIntyre, Alasdair. “Relativism, Power, and Philosophy.” *Proceedings and Addresses of the American Philosophical Association* 59, no. 1 (1985): 5–22.

MacIntyre, Alasdair. *The Unconscious: A Conceptual Analysis*. Revised edition. Routledge, 2004.

MacIntyre, Alasdair. *Three Rival Versions of Moral Enquiry*. Notre Dame UP, 1994.

MacIntyre, Alasdair. *Utilitarianism and the Cost-Benefit Analysis: An Essay on the Relevance of Moral Philosophy to Bureaucratic Theory*. By Kenneth Sayre. University of Notre Dame Press, 1977.

MacIntyre, Alasdair. *Whose Justice? Which Rationality?* Notre Dame UP, 1988.

Mahfuz, Md Lawha. "Understanding the Division Between Analytic and Continental Philosophy." 2024. https://ssrn.com/abstract=5032610.

Mandeville, Bernard. *The Fable of the Bees or Private Vices, Publick Benefits*. Edited by F. B. Kaye. Clarendon Press, 1924. https://oll.libertyfund.org/titles/kaye-the-fable-of-the-bees-or-private-vices-publick-benefits-vol-1.

Mayr, Ernst. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Harvard UP, 1982.

Mill, John Stuart Mill. *On Liberty*. 1859.

Mill, John Stuart Mill. *Utilitarianism*. 1861.

Millgram, Elijah. "Commensurability in Perspective." *Topoi*, no. 21 (2002): 217–26.

Moore, George Edward. *Principia Ethica*. Cambridge UP, 1922.

Mulgan, Tim. *The Demands of Consequentialism*. Oxford UP, 2001.

Murphy, Mark. "The Simple Desire-Fulfillment Theory." *Nous* 33, no. 2 (1999): 247–72.

Nagel, Thomas. *The Possibility of Altruism*. Princeton UP, 1970.

Neumann, John von, and Oskar Morgenstern. *Theory of Games and Economic Behavior*. 3rd ed. Princeton University Press, 1953.

Newman, John Henry. *The Arians of the Fourth Century*. 3rd ed. E. Lumley, 1871.

Nussbaum, Martha. "Human Functioning and Social Justice: In Defense of Aristotelian Essentialism." *Political Theory* 20, no. 2 (1992): 202–46.

Nussbaum, Martha. "Mill Between Aristotle and Bentham." *Daedalus* 133, no. 2 (2004): 60–68.

Nussbaum, Martha. "Practical Syllogisms and Practical Science." In *Aristotle's De Motu Animalium*. Princeton UP, 1978.

Nussbaum, Martha. *The Fragility of Goodness*. Cambridge UP, 1986.

Nussbaum, Martha C. *Women and Human Development: The Capabilities Approach*. Cambridge University Press, 2000.

O'Neill, Onora. "Kant After Virtue." *Inquiry* 26, no. 4 (1983): 387–405.

Paley, William. *The Principles of Moral and Political Philosophy*. Liberty Fund, 2002.

Parfit, Derek. *Reasons and Persons*. Oxford UP, 1984.

Pelikan, Jaroslav. *The Christian Tradition: A History of the Development of Doctrine*. 1: The Emergence of the Catholic Tradition (100-600). Chicago UP, 1971.

Perry, John. "Introduction." In *God, the Good, and Utilitarianism: Perspectives on Peter Singer*, by John Perry. Cambridge UP, 2014.

Perry, John. "Where Did Utilitarianism Come From?" In *God, the Good, and Utilitarianism: Perspectives on Peter Singer*, by John Perry. Cambridge UP, 2014.

Pildes, Richard H., and Elizabeth S. Anderson. "Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics." *Columbia Law Review* 90 (1990): 2121–214.

Pollock, John. *Thinking about Acting: Logical Foundations for Rational Decision Making*. Oxford University Press, 2006.

Postema, Gerald J. "Bentham's Equality-Sensitive Utilitarianism." *Utilitas* 10, no. 2 (1998): 144–58.

Rachels, James. "A Short Introduction to Moral Philosophy." In *The Right Thing to Do*, 4th ed. McGraw-Hill, 2007.

Railton, Peter. *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Cambridge UP, 2003.

Railton, Peter. "How Thinking about Character and Utilitarianism Might Lead to Rethinking the Character of Utilitarianism." *Midwest Studies In Philosophy* 13, no. 1 (1988): 398–416.

Railton, Peter. "Moral Realism." In *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Cambridge UP, 2003.

Rawls, John. *A Theory of Justice*. Revised Edition. Harvard UP, 1999.

Raz, Joseph. "Incommensurability and Agency." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Harvard UP, 1997.

Reeves, Richard. *John Stuart Mill: Victorian Firebrand*. Atlantic Books, 2007.

Reiss, Julian. *Philosophy of Economics: A Contemporary Introduction*. Routledge, 2013.

Richardson, Henry S. *Practical Reasoning About Final Ends*. Cambridge UP, 1994.

Rorty, Amélie Oksenberg, ed. *Essays on Aristotle's Ethics*. University of California Press, 1980.

Rosati, Connie S. “Persons, Perspectives, and Full-Information Accounts of the Good.” *Ethics* 105, no. 2 (1995): 296–325. https://doi.org/10.1086/293703.

Scarre, Geoffrey. *Utilitarianism*. The Problems of Philosophy. Routledge, 1996.

Scheffler, Samuel. “Review: After Virtue.” *The Philosophical Review* 92, no. 3 (1983): 443–47.

Schick, Frederic. *Having Reasons*. Princeton UP, 1984.

Schneewind, J. B. “Moral Crisis and the History of Ethics.” *Midwest Studies in Philosophy* 8 (1983): 525–39.

Schroeder, Tim. “Desire.” In *The Stanford Encyclopedia of Philosophy*, Summer 2020, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University, 2020. https://plato.stanford.edu/archives/sum2020/entries/desire/.

Schultz, Barton. “Henry Sidgwick.” In *The Stanford Encyclopedia of Philosophy*. 2023.

Scruton, Roger. *A Short History of Modern Philosophy: Drom Descartes to Wittgenstein*. 2nd ed. Routledge, 1995.

Sen, Amartya. *Development as Freedom*. Oxford University Press, 1999.

Sen, Amartya. “Internal Consistency of Choice.” *Econometrica* 61, no. 3 (1993): 495–521.

Sen, Amartya. *Rationality and Freedom*. Harvard UP, 2002.

Shapin, Steven. *The Scientific Revolution*. Chicago UP, 1998.

Shaver, Robert. “Utilitarianism: Bentham and Rashdall.” In *The Oxford Handbook of the History of Ethics*. Oxford UP, 2013.

Shields, Christopher. *Aristotle*. 2nd ed. Routledge Philosophers. Routledge, 2014.

Sidgwick, Henry. *Outlines of the History of Ethics for English Readers*. 5th ed. Hackett, 1988.

Sidgwick, Henry. *The Methods of Ethics*. Macmillan, 1874.

Sidgwick, Henry. *The Methods of Ethics*. 2nd ed. Macmillan, 1877.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. Hackett Publishing Company, 1981.

Singer, Peter. “Ethics and Intuitions.” *The Journal of Ethics* 9, nos. 3–4 (2005): 331–52.

Singer, Peter. *Practical Ethics*. 3rd ed. Cambridge UP, 2011.

Singer, Peter. *Rethinking Life and Death: The Collapse of Our Traditional Ethics*. Oxford UP, 1995.

Singer, Peter, and Katarzyna De Lazari-Radek. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford UP, 2014.

Slovic, Paul, and Amos Tversky. "Who Accepts Savage's Axiom?" *Behavioral Science* 19, no. 6 (1974): 368–73. https://doi.org/10.1002/bs.3830190603.

Smart, J. J. C. "An Outline of a System of Utilitarian Ethics." In *Utilitarianism: For & Against*. Cambridge UP, 1973.

Smart, J. J. C., and Bernard Williams. *Utilitarianism: For and Against*. Cambridge UP, 1973.

Sobel, David. "Full-Information Accounts of Well-Being." *Ethics* 104, no. 4 (1994): 784–810. https://doi.org/10.1086/293655.

Sober, Elliott. *Philosophy of Biology*. 2nd ed. Dimensions of Philosophy. Westview Press, 2000.

Steele, Katie, and H. Orri Stefánsson. "Decision Theory." In *The Stanford Encyclopedia of Philosophy*, Fall 2025, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, 2025. https://plato.stanford.edu/archives/fall2025/entries/decision-theory/.

Stocker, Michael. "Commensuration and Public Reason." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Harvard UP, 1997.

Sumner, L. W. *Welfare, Happiness, and Ethics*. Clarendon Press, 1996. https://doi.org/10.1093/acprof:oso/9780198238782.001.0001.

Superson, Anita M. "Deformed Desires and Informed Desire Tests." *Hypatia: A Journal of Feminist Philosophy* 20, no. 4 (2005): 109–26.

Volf, Miroslav, Matthew Croasmun, and Ryan McAnnally-Linz. *Life Worth Living: A Guide to What Matters Most*. The Open Field, 2023.

Vorenkamp, Dirck. "Another Look at Utilitarianism in Mo-Tzu's Thought." *Journal of Chinese Philosophy* 19, no. 4 (1992): 423–43.

Waller, Bruce N. "The Virtues of Contemporary Emotivism." *Erkenntnis* 25, no. 1 (1986): 61–75.

Warburton, Nigel, and David Edmonds. *Philosophy Bites*. Martha Nussbaum on the Value of the Humanities. Aired December 24, 2010.

Watson, David. *Mood and Temperament*. The Guilford Press, 2000.

Weber, Max. "Science as a Vocation." In *The Vocation Lectures*, edited by David Owen and Tracy B. Strong, translated by Rodney Livingstone. Hackett, 2004.

Weijers, Dan. "Nozick's Experience Machine Is Dead, Long Live the Experience Machine!" *Philosophical Psychology* 27, no. 4 (2014): 513–35.

Wiggins, David. “Is Incommensurability Vagueness?” In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Harvard UP, 1997.

Williams, Bernard. “A Critique of Utilitarianism.” In *Utilitarianism: For & Against*. Cambridge UP, 1973.

Williams, Bernard. “Internal and External Reasons.” In *Rational Action: Studies in Philosophy and Social Science*, edited by Ross Harrison. Cambridge UP, 1979.

Wolterstorff, Nicholas. “Locke’s Philosophy of Religion.” In *The Cambridge Companion to Locke*, edited by Vere Chappell. Cambridge UP, 1994.

Yack, Bernard. “Community and Conflict in Aristotle’s Political Philosophy.” *The Review of Politics* 47, no. 1 (1985): 92–112.

Zagzebski, Linda Trinkaus. *The Two Greatest Ideas: How Our Grasp of the Universe and Our Minds Changed Everything*. Princeton UP, 2021.