Principles for Realists:

Political Justification and Genealogical Explanation

by

Brooke Burns

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Philosophy

Dissertation Committee:

Colin Koopman, Chair

Nicolae Morar, Core Member

Erin McKenna, Core Member

Burke Hendrix, Institutional Representative

University of Oregon

Spring 2026

© 2026 Brooke Burns

This work is openly licensed via CC BY 4.0



DISSERTATION ABSTRACT

Brooke Burns

Doctor of Philosophy in Philosophy

Title: Principles for Realists: Political Justification and Genealogical Explanation

An ostensibly intractable divide exists between moral idealists, who treat political principles as applied morality, and political realists, who question whether morality can play a role in politics at all. My dissertation deflates this opposition by arguing for morally austere but politically ambitious principles that nevertheless take seriously the ethical dimensions of political life. My dissertation has two aims. First, I develop a framework for a political justification of thick ethical concepts under realistic conditions. I draw on the work of Elizabeth Anderson, Judith Shklar, Bernard Williams, and Iris Marion Young, all of whom are united in their commitment to a liberal political realism. Second, I argue that genealogical explanations, and specifically empirically adequate genealogies, provide the descriptive resources required for politically realist justifications. I consider the idealized functionalist models of Miranda Fricker and Matthieu Queloz as well as Williams's vindicatory genealogy, before arguing that Michel Foucault's genealogical method is best suited to a politically realist inquiry. I then mobilize my politically realist genealogy to consider in detail the thick ethical concept of dignity in United States prisons. I argue that dignity serves a dual function: it both constrains the mechanistic application of rules and establishes standards of punishment. To make a demand in the name of dignity is to identify forms of treatment for which there is no penological justification.

ACKNOWLEDGEMENTS

Writing demands a lot of time alone. It equally requires good conversations, keen and gracious readers, and words of encouragement when doubt sets in.

No one has contributed more to my education in philosophy than my advisor, Colin Koopman. Colin is a paragon philosopher, teacher, and mentor. He has read every draft of this dissertation. Through our conversations, he helped work out its ideas before and between drafts. He taught me how to be a careful and gracious reader of philosophy. His encouragement kept the doubt at bay. The best way I can express my gratitude is by striving to become the kind of philosopher and mentor to my own students that Colin has been for me.

I am also grateful to my committee members — Erin McKenna, Nicolae Morar, and Burke Hendrix — for reading my dissertation and supporting my project through its development. I am especially thankful to Erin and Nicolae for the many seminars in which they taught me how to think like a pluralist.

Thank you to Gonzalo Bustamante Moya for forming the supportive bedrock of the philosophy department, for being both a friend and colleague, and for always making the time. To John Montani, for our endless (and endlessly enjoyable) conversations about philosophy, teaching, and much more. To Jensen Wainwright and Mariam Nadeem, especially in my final year of writing. In the grey Eugene winters, they always brightened my day.

I am grateful to the Oregon Humanities Center for providing me with financial support, a term off from teaching, and an office to write in during Spring 2025. Finally, I thank all the faculty, staff, and graduate students with whom I spent many hours in Susan Campbell Hall over the years. I am grateful to have loved writing my dissertation and to have done so in the philosophy department at the University of Oregon.

For my mother, Mary

Table of Contents

DISSERTATION ABSTRACT **3**

ACKNOWLEDGEMENTS **4**

I. INTRODUCTION: THE PROBLEM **8**

1. Idealism and Realism: Varieties of Wishful Thinking 8
2. Realistic Inquiry 10
3. Dignity for Realists 12
4. Overview 14

II. BETWEEN THE IDEAL AND THE REAL IN RAWLS, SHKLAR, AND WILLIAMS **17**

1. Ideal Theory: The Strong and the Practicable 19
2. Ideal Theory as Idealized Justification 23
3. Liberal Political Realism 29
4. Putting Cruelty First 38
5. Thickening Cruelty 45
6. Critical Realism 49

III. FEMINIST POLITICAL REALISM IN ANDERSON AND YOUNG **54**

1. Pluralist Values and Group Relations 56
2. Oppression 62
3. Inclusive Political Justification 65
4. Broad and Narrow Internal Exclusions 68
5. Ideals of Individual and Collective Empowerment 75
6. From Social Critique to Genealogical Critique 77

IV. ON WHAT EXPLAINS: STORIES, MODELS, AND PARADIGMS **79**

1. Genealogy and the State of Nature 80
2. Three Uses of History in "Truth and Truthfulness" 87
3. Idealized Functionalist Models 94
4. Realistic Explanations 102
5. Explanatory Choices 107
6. Realism and Contextual Values 115

V. FOUCAULT'S *AUTOPSY* **119**

1. Seeing For Oneself 121
2. Foucault's Meta-History 125
3. Rupturing the Past: The Elements of Archaeology 129

4. Is Savoir Sufficient? 138
5. Emergence and Descent: The Elements of Genealogy 145
6. Value Judgements and Explanatory Choice 148
7. Fabricating Reality 153

VI. A POLITICALLY REALIST GENEALOGY OF DIGNITY 155

1. Politically Realist Genealogies 158
2. The Birth of Dignity 164
3. Type I: Proportionality Dignity Violations ca. 1957 168
4. Type II: Culpability Dignity Violations ca. 1971-2004 172
5. Type III: Treatment Dignity Violations ca. 1967-1974 177
6. The Emerging Rights of the Confined 182
7. Putting Dignity First 190

BIBLIOGRAPHY 193

I.

INTRODUCTION: THE PROBLEM

"You've seen trees by a raging winter torrent, how many sway with the flood and salvage every twig, but not the stubborn—they're ripped out, roots and all. Bend or break. The same when a man is sailing: haul your sheets too taut, never give an inch, you'll capsize, and go the rest of the voyage keel up and the rowing-benches under. Oh give way. Relax your anger—change!"
—Sophocles, *Antigone*

1. Idealism and Realism: Varieties of Wishful Thinking

Two dominant views on the relationship between ethics and politics persist in philosophy. On the one hand, the moral idealists, who in their unwavering commitment to universal principles, grow contemptuous of political life. Confronted with the fact of human folly and the inertia of pragmatic reform, the moral idealist demands from the world more than it can give or else withdraws from the world, if only to preserve the purity of their ideals. On the other hand, the political realists, who having grown disenchanted with the moral universe, view morality as either a form of strategic weakness or ineffective in guiding politics. Or perhaps in the course of their lives, they have seen the harms that moralism itself creates by those who have been morally injured and who have in turn exerted injury on others in the form of vitriol and resentment. These views are united by their skepticism about whether there is really such a thing as morality in politics at all; if morality does exist in politics, they wonder whether it is consequential in our actions, whether we can rely on it when things get tough, and whether we can maintain our integrity when conflicts and compromise inevitably arise. Moral idealism and political realism are but two sides of the same coin. The desire for purity encourages skepticism about morality's hold on the world while the disenchantment of realism strips political life of ethical significance altogether.

This problem is not unique to our time. In Sophocles' *Antigone,* the play's namesake defies Creon's laws out of an unwavering commitment to the divine laws. She scorns those who fail to assist in the task of burying her brother, particularly her sister Ismene, who urges Antigone against breaking the state laws and bringing even more tragedy upon the family. Antigone is

unsympathetic to Ismene's conflicting commitments. Creon, the king Antigone defies, is just as obstinate in his commitment to defend his authority and the interests of the state against Antigone. Despite his son Haemon's pleas to his father to be more understanding, Creon is inexorable. Antigone must be punished at any cost. And the cost is high indeed. Antigone commits suicide in defiance and what ensues is a fate worse than death for Creon. Sophocles' *Antigone* expresses a hard lesson in the pitfalls of overly principled commitment, be it to the authority of moral laws or of state legality. It also demonstrates that these two worldviews are united by a shared obstinacy, which they take to be a mark of integrity, but which ultimately wreaks tragedy upon themselves and those around them.

In our contemporary moment, both moral idealism and political realism are varieties of wishful thinking, in which the desire for something to be true leads us to believe that it indeed is true. Wishful thinking is ultimately a refusal to listen to the evidence that disproves our desires. Moral idealists desire exemplary people and institutions and accordingly constructs their principled commitments on the basis of that aspiration. Political realists assert that humans are weak in integrity and full of faults and resigns themselves to self-defeating principles in light of that fact. The realist's wishful thinking stems less from the conscious desire that humans are morally flawed than from a disposition of resignation. Realists hope for others to behave ethically and when they fail to live up to those expectations realists feel let down. Rather than feel the sting of this disappointment, they construct principles on the bedrock of defensive pessimism, misanthropy, and cynicism.

How can we find our way out of this dichotomy? How can we avoid our own tendencies toward idealistic or cynical wishful thinking? One response lies in having the courage to be disillusioned, which is another name of coming to realize that one's deepest held views of the world are mistaken. Disillusionment need not lead us down a path of despair. Indeed, it can strengthen our commitments by making them more responsive to how the world actually is and thereby more effectively shape how we believe the world ought to be. Such an engagement requires a mode of inquiry that resists wishful thinking and that brings idealism and realism into closer relation.

2. Realistic Inquiry

I argue for a realistic approach to ethical concepts in theorizing how such concepts ought to guide political action. This requires, I suggest, a realistic explanation of those concepts, understood as one that is empirically adequate. In this sense, it is more appropriate to describe the "political realist" discussed above as a *denier* of ethics, insofar as she denies the significance of the ethical in political life. There is nothing problematic about being realistic provided that realism does not dogmatically or uncritically reject normative ideals. One way a political realist of the denialist persuasion can be uncritical is by accepting deeply fraught practices without first questioning their authority. This is the general concern idealists raise about ethical theories that rely on practice-dependent ideals, insofar as such practices often reflect received views that are shrouded in bias, prejudice, and other negative attitudes, which ethics seeks to dispel. This worry, however, overlooks how a practice-dependent inquiry can itself be a site of critical reflection.

Justification begins with shared evaluative judgements.[1] These judgements define what counts as a reason, and they are typically articulated through what Bernard Williams calls "thick ethical concepts."[2] Thick ethical concepts are descriptive, or "world-guided," insofar as particular facts must obtain for their application. They are also evaluative, or "world-guiding," insofar as they provide reasons for acting in specific ways. Thick ethical concepts evoke an image or images of the correct actions and conditions that constitute an appropriate expression of that concept, along with a valuation that expresses them in terms of a complex set of emotions and dispositions. When I hear the word "courage," for example, I imagine situations involving risk and fear and someone acting according to what is right despite those risks. I also express a set of complex emotions towards the courageous person or action, such as admiration and respect.

What is important about thick ethical concepts for my account is that they comprise the material of our justifications. When we disagree over thick ethical concepts, we challenge the factual basis of another's judgements by offering reasons for why a particular action does not fall under the concept in question. We can also challenge the authority of our judgements by offering reasons for why it no longer makes sense to value a given ethical concept. Likewise, we can

[1] See Elizabeth Anderson, *Value in Ethics and Economics* (Harvard University Press, 1993), 104-112.

[2] Bernard Williams, "Knowledge, Science, Convergence," in *Ethics and the Limits of Philosophy* (Harvard University Pres, 1985), 132-155.

vindicate their authority by giving reasons for why we should care about a particular concept. Often in practices of justification, we disagree on both the facts and the evaluation: I challenge your interpretation of a thick ethical concept and give you reasons for why my interpretation is not only the correct application of that concept, but also why it makes more sense for us to care about the concept in that way. When we inquire into the factual basis and authority of our ethical practices, we are wondering whether they make sense as both intelligible and appropriate expressions of those practices and also whether we have good reasons for endorsing them. One way we do this is by subjecting them to critical scrutiny. For this reason, I argue that a genealogical inquiry is an appropriate way of making sense of thick ethical concepts.

There are those who believe that a genealogical method can, on its own, demonstrate the factual basis of a practice and in turn generate sufficient normative force to affect the space of reasons. Friedrich Nietzsche, for example, uses genealogy to demonstrate the factual basis of morality as grounded in *ressentiment* against noble and powerful people, a realization which is itself reason-giving for us to condemn Christian morality. Contrasting with Nietzsche's subversive genealogy, Williams's vindicatory genealogy demonstrates the universal need for the virtues of truth, which in turn gives us reasons to engage in practices of truthfulness.

A method that blends explanation and justification is not without controversy. I am skeptical of hybrid accounts and argue that, in keeping the two components separate, we can yield the same (if not better) explanatory and justificatory power without nearly as many philosophical complications. Such a genealogical approach is most closely tied to the work of Michel Foucault, whose genealogies generate neither undermining nor vindicatory reasons for our endorsement of practices. But if a genealogy of this kind is not itself reason-giving, how does genealogy provide critical insight into our practices?

I argue that genealogies enable us to grasp the descriptive nature of thick ethical concepts by first demonstrating that the concept is indeed a *thick* ethical concept. It is often the case that we are unaware of our ethical concepts as world-guided and instead take our intuitive judgements of those concepts as the appropriate and universal application of them. When others fail to express those concepts in ways we would deem appropriate, we might judge or deride them rather than inquire into the factual content of the concept itself, largely because we do not even see that it is grounded in the facts. Second, genealogy gives an account that makes sense of both the descriptive nature of a thick ethical concept and by apprehending both the conditions under

which the concept can be expressed and the concrete practices that express those concepts. Third, genealogy gives an account of the evaluative nature of a thick ethical concept, by demonstrating the associated attitudes expressed in those practices and the web of further thick ethical and *counter* ethical concepts in which it exists. Only once we have a clear grasp of a thick ethical concept, can we go on to subject our practices to critical scrutiny in a practice of justification. Genealogy is not the only critical strategy realists might draw on, but it is a very promising one.

3. Dignity for Realists

Consider dignity, one of our thickest ethical concepts. In the philosophical literature, dignity is in disarray. Should we define a set of necessary and sufficient properties of dignity or determine what grounds dignity? Is dignity a matter of character, value, or status? Is it something someone possesses qua an individual, or as a member of a particular class? Is it inherent, or can it be bestowed and revoked? With the dignitarian's inveterate quest to establish a sound and moral account, the dignity denier becomes even more suspicious. Notwithstanding the denier's suspicion, the two positions have more in common than they may initially suppose. Together they exemplify the broader confrontation between idealism and realism. What unites them is that they both expect a great deal from dignity. When dignity fails to meet these expectations, dignitarians tend to fault individuals or institutions for failing to live up to its demands, while deniers fault the concept itself and conclude we are better off without it. But are we really? I argue we are not or, at the very least, that we should first clarify the historical conditions, practices, and valuations associated with dignity before issuing judgement. The problem, I argue, is not with dignity, but with our impoverished understanding of its thick content.

How can genealogy provide a fuller understanding of dignity in ways that other philosophical approaches have not? First, by remaining agnostic about a *theory* of dignity. This is because thick ethical concepts are characteristically pluralistic. What might count as an intelligible expression of an ethical concept in one context may not obtain in another. The practices I express when I treat my students with dignity will be different from the practices a corrections officer expresses when he treats a prisoner with dignity. There may be a shared kernel underlying these practices, but that is not for genealogy to prove.

Second, genealogy can demonstrate *that* dignity is a thick ethical concept. Some concepts are readily accepted as thick insofar as they are clearly born of a particular historical moment and

bear that moment's descriptive and evaluative weight, such as "chastity" or "chivalry." Others are less obviously historical, in part because philosophical discourse often abstracts them from practice through highly idealized theories of the concept's essential nature or necessary and sufficient conditions. With respect to dignity, a genealogical inquiry can provide an account of the historical moment at which it emerged as a thick ethical concept.

Third, genealogy can help us clarify the descriptive content of dignity. It provides an account not only of its emergence, but also of the conditions under which it emerged, and the practices associated with it. This third use of history underscores the diagnostic and critical significance of genealogy for thick ethical concepts. Consider, for example, the dignitarian who holds that treating someone with dignity requires regarding that person as a moral equal. The practices associated with this understanding include addressing the person by name, listening when she speaks, and refraining from contemptuous attitudes. Now consider removing this agent from a context in which egalitarian relations largely obtain, such as the seminar room, and placing him in the role of a corrections officer in a prison. Would the content or expression of dignity change? Dignitarians will often insist that the concept and function of dignity would not change. If so, is the corrections officer morally obligated to act on this universal notion of dignity even when the normative avenues for expressing it do not obtain? For now, I raise this question only to illustrate how genealogy can illuminate the gap between our intuitive evaluations of a concept and its socially intelligible expressions.

Fourth, genealogy can provide a descriptive account of the evaluative content of dignity by situating it within a web of other thick ethical and counter-ethical concepts, such as respect and autonomy, cruelty and humiliation. It can also bring into view first- and second-order attitudes toward dignity, which is to say, the standards a group sets for dignity and the processes by which it judges the extent to which existing practices meet those standards. Once again, genealogy functions as a diagnostic rather than an evaluative tool, even when the content it diagnoses it itself evaluative.

If, after subjecting dignity to critical scrutiny, we come to believe that our personal valuations of dignity exceed what is presently intelligible, we may grow disillusioned with dignity. My hope, however, is that such disillusionment can enable a more effective interrogation of the received understanding of dignity so that its practices better live up to our personal valuations.

4. Overview

The majority of this dissertation engages in debates in political philosophy and genealogy and develops what I call a *politically realist genealogy*. The dissertation is divided into two parts. The first part develops a framework for a realistic justification by articulating two complementary methodological commitments in liberal political philosophy. First, I argue that normative ideals should be constructed in light of what values have been locally secured, rather than through idealized accounts of how the world ought to be. Second, I argue that political philosophy must attend to how justificatory practices actually function in contemporary political life. Contrary to some worries, a more realistic approach to politics need not resign itself to the status quo. Indeed, it must make a conscious attempt to improve the force of our normative ideals and the procedures through which they are justified. In the second part of the dissertation, I argue that genealogical explanations supply the descriptive resources required for a politically realist justification. Genealogies trace the emergence and development of thick ethical concepts and bring their descriptive and evaluative dimensions into view. Not all genealogies are well suited to this task, however. To help us adjudicate between competing explanatory methods, I argue we require set of criteria for judgment. Finally, in Chapter Six, I mobilize my politically realist genealogy to consider in detail the thick ethical concept of dignity in contexts of United States prisons.

In Chapter Two, I set out what I take to be the point of political philosophy: to guide our actions and to make our world a better place than it presently is. As much of a platitude as this may sound, the premise upon which my dissertation rests is not without controversy. I begin by devoting attention to strong utopians, who care less about the practical applicability of political philosophy and more about constructing a sound understanding of ideals and justificatory models of people being better than they are capable of being. I disagree with this conception of political philosophy and turn first to the practicable utopianism of John Rawls and then to the political realism of Judith Shklar and Williams to argue that the point of political philosophy is to construct ideals with normative force. The difference between Rawls and realism is one of degree rather than of kind. Though they agree on the point of political philosophy they disagree on its scope. More specifically, they disagree over how much we can expect from institutions and individuals to behave morally in politics. I attribute the nature of this disagreement to how each view conceives of the relation between the political and the moral spheres in justifying the

rightful exercise of power. Rawls conceives of politics as applied morality. Shklar and Williams, by contrast, disambiguate the relation between the moral and political and incorporate a historical sense into their liberal normative theories. Given that liberalism is a comparatively recent political arrangement, they argue that it does not make much sense to claim that rightful power and liberal aspirations to social justice are one and the same. They argue that what liberalism offers above and beyond other political arrangements is an amelioration of cruelty. The ideals we ought to vindicate, therefore, are not the ones that lead to greater justice, but that prevent the worst forms of cruelty.

In Chapter Three, I draw on the work of Elizabeth Anderson and Iris Young to develop a much-needed account of a feminist political realism. I argue that such feminist realism shares three methodological components: a pluralist metaethical framework, a negative aim of ameliorating oppression, and a positive ideal of individual and collective empowerment. At the metaethical level, feminist realism is committed to value pluralism. Our valuations of persons and practices, and their appropriate modes of expression, are socially situated and context-sensitive. Given this pluralism, we are bound to disagree over what ought to be valued and appropriate modes of expressing our valuations. For this reason, feminist realists are reluctant to construct ideal normative principles and instead argue for ideal procedures through which contested values can be critically assessed. In ideal circumstances, justificatory procedures manifest both the negative and positive aim of feminist realism by enabling meaningful participation and constraining the legitimation of oppressive values. Yet existing practices fall woefully short of the idealized accounts we find in Rawls and other utopians. In response, I argue for an ideal of individual and collective empowerment, by which oppressed individuals recognize and critically assess the values that structure their lives. This ideal serves as a first step towards meaningful social critique.

In the second half of the dissertation, I consider philosophy's role in strengthening social critique. I argue that genealogy can aid in a process of collective criticism by illuminating the descriptive and evaluative dimensions of our practices. Only once we have made sense of those practices can we subject them to critical scrutiny. Not all genealogies are suitable for aiding a politically realist justification. This is because some genealogies rely more heavily on idealizations of human conduct while others draw on more empirical methods. For the purposes of my project, I argue that we ought to prefer genealogical explanations that are empirically

adequate, which I understand as the degree to which an explanation matches our observations of how humans actually behave. The criteria I provide for assessing empirical adequacy are determinacy, impartiality, contingency, and regularity.

In Chapter Four, I first consider genealogical explanation in contemporary analytic philosophy. Recent interpretations of Williams's vindicatory genealogy have cast his method as primarily fictional and expressive of practical necessity before it is historical and expressive of contingency. I contest this interpretation of Williams's method and argue that his genealogies reveal the indispensability of history for explaining why and how we have come to adopt ethical practices. I then show how the lack of historical sense in idealized functionalist models in analytic genealogy are not suitable for social critique because they do not meet the four criteria for empirical adequacy. By contrast, Williams's genealogy satisfies these criteria.

Chapter Five then turns to Foucault's genealogical method, which I argue also satisfies the criteria for empirical adequacy. Considering both Williams and Foucault meet these criteria, is choosing between the two methods ultimately a matter of preference? I argue it is not. To adjudicate between empirically adequate approaches, I introduce two further criteria. The first, *freedom*, favors methods that adequately represent human capacities for transformation. The second, *complexity of interaction*, favors methods that represent the dynamic interaction among explanatory elements, rather than reducing them to a causally determinate structure. With these two additional criteria, I argue we should prefer Foucault's genealogical method.

Finally, in Chapter Six, I provide a politically realist genealogy of dignity as a thick ethical concept. In my politically realist genealogy, I argue that dignity has historically functioned as a morally austere but politically ambitious threshold concept that has constrained the two poles of legalism. On the one hand, dignity curtails the overly strict and mechanistic application of rules and ensures the individual is seen and treated as an individual; on the other hand, it establishes clear standards and constrains the retributive cruelty of the state. To make a demand in the name of dignity, on my view, is to identify forms of treatment for which there are no justifications within a practice. To the extent that dignity serves a role, genealogy shows that it serves a very important one. Whether we think it should continue to serve this role is up to us to determine.

II.

BETWEEN THE IDEAL AND THE REAL

IN RAWLS, SHKLAR, AND WILLIAMS

In her 1957 book *After Utopia: The Decline of Political Faith* Judith Shklar explored the diminishing rational optimism that once defined the Enlightenment. According to Shklar, talk of justice had become "intellectually hazardous," hopes for practical reason had given way to fatalism, and political thought had been rendered inert.[3] Shklar did not offer a solution in *After Utopia,* in part because none was present at the time. What she did provide, however, was a rejoinder to the complacent cynicism that had usurped utopianism. We pursue ideals like justice, she argued, not because we naively believe in the possibility of perfect social harmony. Without ideals we lose "the very possibility of social knowledge and amelioration."[4] Already a stark realist, Shklar called for us to pursue ideals simply because it is better to do so than to not.

In 1971, John Rawls issued a reappraisal for political philosophy with *A Theory of Justice* and subsequently gave rise to a renewed interest in the normative force of ideal principles, a force still present but waning in contemporary political philosophy. From this recurring disillusionment with the role of the ethical in political life, a purported divide has appeared.[5] On the one side of the divide are the ideal theorists, with Rawls often serving as the paradigm; on the other, the political realists, a label I attribute to Shklar and Bernard Williams. Although significant differences exist between Rawls, Shklar, and Williams, I argue the literature tends to

---

[3] Judith Shklar, *After Utopia: The Decline of Political Faith* (Princeton University Press, 1957), 272.

[4] Shklar, *After Utopia,* 18.

[5] A third approach to these debates is nonideal theory, a method in political philosophy that I view as closely related to Shklar and Williams, though I concede that some view nonideal theory as more akin to applied ideal theory. Likewise, there is disagreement within political realism regarding whether it serves to primarily critique, correct, or replace liberal political philosophy. See, for example, William Galston, "Realism in Political Theory," *European Journal of Political Theory* 9, no.4 (2010), 385-411; for a comprehensive overview of the debate see Laura Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map," *Philosophy Compass* 7, no.9 (2012), 654-664). For now, I focus on realism as a positive corrective to liberal political philosophy as opposed to a non-liberal alternative. This makes my reading more aligned with "nonideal" theorists such as Elizabeth Anderson and Charles Mills than other political realists such as Raymond Geuss.

overemphasize these differences. What we lose in doing so, are the deeply ethical dimensions of Shklar and Williams's political realism.

Not all realists agree on the role of normative ideals within politics. Whether one is sufficiently convinced by the possibility of a normative project depends on the kind of realism she aligns herself with and the threshold of conflict she is willing to accept in politics. If, for example, she follows Machiavelli and accepts the world as a bleak place of unrestrained violence, she may see the necessary backdrop of politics as a conflict of cruelty. If, however, she refuses to accept such a world, she will instead read *The Prince* as Shklar says Montaigne did, that is, "as a victim of princely cruelty might read it."[6] For Shklar, putting cruelty first is as much a response to ideal theory as it is to "the new science of politics" we find in Machiavellian realism.[7] Shklar and Williams acknowledge that conflicts of plurality—of loyalties, of obligations, of desires, or of ethical responsibilities—may always be present in an uncertain world. Given this plurality, our normative projects should diminish cruelty.

In this chapter, I explicate, refine, and justify, the normative commitments of Shklar and Williams. First, I provide an overview of ideal theory in political philosophy, by distinguishing between the *strong utopianism* of G.A. Cohen and David Estlund against the *practicable utopianism* of Rawls. I examine the realism of Shklar and Williams and demonstrate their shared commitment to liberal political values. I then conclude by considering Shklar's morally austere but politically ambitious "cruelty principle."

One notably absent figure from this account is Raymond Geuss, who shares many significant affinities with Shklar and Williams. Like them, Geuss is critical of Rawlsian ideal theory. He also argues for a historically situated approach that asks not only what political philosophy ought ideally to do, but the role it has actually played in political life.[8] Although his work is indispensable for understanding the heterogeneity of political realism, my focus here is on a specifically liberal political realism. Geuss, by contrast, does not aim to reconstruct normative ideals that, while realist, retain a distinctively liberal character. Indeed, one aim of Chapters Two and Three is to construct a form of political realism in liberal figures not typically

[6] Judith Shklar, *Ordinary Vices*, 11.

[7] Shklar, *Ordinary Vices*, 11.

[8] Raymond Geuss, *Philosophy and Real Politics* (Princeton University Press, 2008), 37.

read as realists, most notably Shklar in this chapter, and Elizabeth Anderson and Iris Marion Young in the following chapter. In so doing, I excavate a liberal political realism that aims to construct historically situated and normatively effective ideals.

1. Ideal Theory: The Strong and the Practicable

Ideal theory has been the prevailing method in normative and political philosophy at least since Rawls coined the term, though its methodological dimensions are seldom understood because so rarely interrogated. This is true of those who claim it is useless, naïve, and ideological, as well as those who presume it is "the only basis for the systematic grasp" of our non-ideal world.[9] For those who adhere to ideal theory, the presumption goes as follows: even if a philosopher is concerned with, the systematic injustices of everyday life, she must first hold an ideal of justice in view. And if a political philosopher considers an ethical collective life through ideals, ideals which posit a world better than the one we currently inhabit, then she must aim to realize that ideal through principles. I call this *the ideal as evaluative standard*.[10] Rawls's theory goes further than this first sense of ideal theory. His approach constructs principles through a model of justification known as the original position. I call this *the ideal as model*. In this section, I capture both dimensions of Rawls's ideal theory with my aim being a much-needed analysis of Rawls's ideal as model, an often mischaracterized and thus misunderstood methodological dimension of Rawls's ideal theory. By way of this distinction, I will show that even when disagreements between ideal and non-ideal theory run deep, they are not as intractable as they initially appear.

To begin, ideal theory sometimes implements the ideal as an evaluative standard. In these circumstances, we say: 'anything less than P is not ideal'. With reference to a theory of justice, we say: 'anything less than P is unjust'. An evaluative standard expresses a political principle as an ideal theory and contains multiple senses. For example, an ideal may indicate a Platonic form, a utopia, or that which is ultimately impossible to achieve. I refer to this form of ideal as evaluative standard as *strong utopianism*. Alternatively, when one uses an evaluative standard

[9] John Rawls, *A Theory of Justice* (Harvard University Press, 1971), 8.

[10] See Charles Mills, "Ideal Theory as Ideology," *Hypatia: A Journal of Feminist Philosophy* 20, no. 3 (2005), 165-183; Zofia Stemplowska and Adam Swift, "Ideal and Nonideal Theory," *The Oxford Handbook of Political Philosophy,* ed. David Estlund (Oxford University Press, 2013), 373-390.

within the realm of what is practicable, as we find in Rawls, I refer to this as *practicable utopianism*. For both grades of evaluative ideal theorists, principles express a social ideal and so guide social reform.[11] A social ideal, according to Rawls, is deeply rooted in a concept of society and reflects how such a society conceives of the aims and purposes of cooperation.[12] It is worth noting here that ideal theorists often conflate justice with ideal theory itself since it is seen as the first virtue of social institutions. And yet, as Rawls himself notes, the principles of justice are but one part, even if they are the most important part, of that ideal.[13]

Rawls's ideal principles of justice give ideal theory its institutional form and guide the action of social conduct.[14] When rules are impartially and consistently administered, we have what Rawls calls formal justice, understood as strict obedience to a principle. Formal justice protects citizens from the arbitrary application of rules, though, from the standpoint of the least advantaged, it merely provides impartiality and consistency from an otherwise unjust institution. Rawls is inclined to believe, however, that when formal justice is upheld, substantive justice will also exist. Where there is a desire to follow rules impartially and consistently, we also find a motivation to "recognize the rights and liberties of others and to share fairly in the benefits and burdens of social cooperation."[15]

Rawls holds that a just institution is only realized through the expression of a public system of rules. An institution exists as an abstract object, that specifies certain modes of conduct as permissible or forbidden according to the two principles.[16] These principles are then realized "in the thought and conduct of certain persons at a certain time and place."[17] A substantively just

[11] For comprehensive overviews of the heterogenous literature, see Laura Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map," and Stemplowska and Smith, "Ideal and Nonideal Theory."

[12] Rawls, *A Theory of Justice*, 9.

[13] Rawls, *A Theory of Justice*, 9.

[14] In *A Theory of Justice*, Rawls addresses two principles of justice. The first principle ensures equal liberties for all while the second requires social and economic inequalities to satisfy two conditions (a) that they are attached to offices and positions open to all under conditions of fair equality of opportunity; and (b) they are to the greatest benefit of the least advantaged (266).

[15] Rawls, *A Theory of Justice*, 52.

[16] Rawls, *A Theory of Justice*, 47-48.

[17] Rawls, *A Theory of Justice*, 48.

institution is thus one in which the two principles of justice are expressed and embodied in the thought and action of individuals, as a public system of impartially applied rules.

Here we arrive at one fundamental disagreement between strong and practicable utopians. Practicable utopianism argues that we only assess the justice of an institution as it is realized. Strong utopians, by contrast, hold that ideal institutions cannot, and perhaps should not, be realized. This is an incommensurable characterization of justice and one that, as I will argue later, represents a far greater gulf between strong and practicable utopianism than between practicable utopianism and political realism. For Rawls, justice cannot occur between "an association of saints agreeing on a common ideal" since it only comes about through practices where there are competing interests and when persons feel entitled to press their rights on each other.[18] Perfect agreement may indeed be theoretically exemplary (if not practically impossible), but it does not express a just society.

Indeed, the disagreement runs even deeper. Strong utopians claim that justice is more than a mere instrument of practical reasoning. Estlund holds that the value of justice is similar to the value of advanced mathematics or theoretical physics.[19] Practicalism would declare this a waste of time, but we of course know it is not. Strong utopianism does not concern itself with practical problems, such as feasibility or applicability, but instead with constructing a "sound understanding of the nature of justice."[20] Similarly, Cohen writes, "the question for political philosophy is not what we should do but what we should think, even when what we should think makes no practical difference."[21] Political philosophy may occasionally stumble upon practicality, but this is not its primary aim.[22]

---

[18] Rawls, *A Theory of Justice*, 112.

[19] David Estlund, *Utopophobia: On the Limits (If Any) of Political Philosophy* (Princeton University Press, 2019).

[20] Estlund, *Utopophobia*, 37.

[21] G.A. Cohen, "Facts and Principles," *Philosophy & Public Affairs* 31, no. 3 (2003), 243.

[22] This sentiment has been echoed by Stemplowska, who argues that the purposeful inclusion of false assumptions can often prove effective for answering a range of questions fact-sensitive theories are unequipped to handle. For example, we may include chimerical assumptions or construe an overly simplified universe in order to assess the relation between two principles. According to Stemplowska, those who hold that normative theory is *solely* a guide for action miss out on its ability to judge states of affairs. See Zofia Stemplowska, "What's Ideal About Ideal Theory," *Social Theory and Practice* 34, no. 3 (2008), 328.

Flowing from the first disagreement, the second concerns the fact-sensitivity of principles. A practicable utopian holds that facts are pertinent to principle construction and imagines an ideal with deficiencies in view. For Rawls, ideals represent what we would be capable of if we were better. Strong utopians, by contrast, imagine a society without such flaws by requiring within its scope "people being better than they will be."[23] Estlund calls this a *non-concessive* view, insofar as it refuses to concede to the facts. A non-concessive view does not assume different facts, it simply ignores them in favor of other-worldly principles.

Cohen is the exemplar of the non-concessive philosopher. For Cohen, there are fundamental principles which are independent of and logically prior to facts. In addition, these principles do not depend on human nature or psychology, nor need they bother with pragmatic concepts of feasibility or compliance. Cohen's argument for this view goes as follows: every fact-sensitive principle (P) reflects a fact-insensitive principle (P′), which in turn justifies why a fact (F) matters to P in the first place.[24] Strong utopianism maintains that there are fundamental and timeless principles that are morally grounding but are themselves ungrounded in nonmoral facts.[25] What justice demands holds strong despite changes in facts such as attitudes, history, and social norms. When we apprehend the nature of justice, our judgment should not be clouded by contingent whims.

Rawls agrees that principles must hold some fundamental value since his theory seeks to be sufficiently general and wide-reaching. But if principles want to guide us within our world, they must equally be minimally fact-sensitive. Conditions under the practicable ideal are thus reasonable and exist in a possible historical condition allowed for by the laws and tendencies of the social world.[26] Justice as fairness, Rawls writes, "probes the limits of the realistically practicable, that is, how far in our world (given its laws and tendencies) a democratic regime can

[23] Estlund, *Utopophobia*, 121.

[24] Cohen, "Facts and Principles," 215.

[25] Estlund, *Utopophobia*, 45.

[26] Rawls, *Justice as Fairness: A Restatement* (Harvard University Press, 2001), 4.

attain complete realization of its appropriate political values."[27] Though Rawls may be utopian he is always, in his own phrase, realistically utopian.[28]

Realizing a practicable ideal only seems possible if everyone were to strictly comply with the public systems of rules. In many ways, ideal theory for Rawls is conterminous with strict compliance.[29] But Rawls does not believe strict compliance will always (or even often) come easily. Many times, we require institutions to implement compliance. Sometimes, these institutions are mere nudges in the right direction and other times they take on a stronger coercive force. Either way, the underlying assumption remains that humans could comply since the ideal principles are minimally constrained by the facts. This is why Rawls is hesitant to presume supererogatory motivations and actions such as heroism, benevolence, mercy, or self-sacrifice.[30] With practicable utopianism we are asked to strictly comply though not beyond the realm of what humans could reasonably accomplish.

2. Ideal Theory as Idealized Justification

Ultimately, Rawls is not entirely satisfied with justice as an evaluative standard since this merely represents our provisional convictions at best. For example, if a person believes that gender discrimination is unjust, and she feels confident in this conviction, she is simply positing a self-evident premise.[31] As such, Rawls places a constraint on justifications by excluding from the start "claims to universal truth, or claims about the essential nature and identity of persons."[32] One must, in other words, provide a political and not metaphysical justification for her considered judgements.

To arrive at the ideal principles of justice as fairness Rawls steps back "to envision our objective from afar."[33] Just as principles are deemed just or unjust only when they are realized in

---

[27] Rawls, *Justice as Fairness: A Restatement*, 13.

[28] Rawls, *Justice as Fairness: A Restatement*, 4.

[29] See Rawls, *A Theory of Justice*, 7-8; *Justice as Fairness*, 13.

[30] Rawls, *A Theory of Justice*, 100.

[31] Rawls, *A Theory of Justice*, 19.

[32] Rawls, "Justice as Fairness: Political Not Metaphysical," *Philosophy & Public Affairs* 14, no. 3 (1985), 223.

[33] Rawls, *A Theory of Justice*, 19.

institutional form, we can only deem ideal principles as just when they are the outcome of a fair procedure. The problem, however, is that we cannot offer an entirely empirical confirmation of pure procedural justice, much in the same way that we cannot offer an empirical confirmation of fundamental physics. Rawls's solution is an idealized situation best understood as a model of fair procedural justice, which is the second sense of ideal theory, the ideal as model.[34] Rawls's original position models the consequences of a fair and rational outcome of a hypothetical social contract instead of modeling, say, humans as they are. Crucially, the model does not presuppose principles of justice as independent criteria for the just result. A pure procedure is one where a system of rules is freely entered into under fair conditions.[35] We know we have reached a fair outcome not because we have defined the outcome in advance, but because the procedure has been fairly carried out.

Rawls first sets out the background circumstances of his model and tries to do so without smuggling in theories of human motivation or of human nature.[36] In this way, an ideal model is built on a set of weak conditions generally shared but strong enough to yield a set of principles.[37] How individuals actually deploy practical reasoning in moral conduct is orthogonal to modeling an initial situation with conditions that we do in fact accept.[38] The first condition defines the circumstances of justice, which Rawls describes as "the normal conditions under which human cooperation is both possible and necessary."[39] For example, claims of distributive shares only take place against a backdrop of moderate scarcity, where individuals share in the confidence that social cooperation makes for a better life than one lived alone. But as social cooperation gives rise to greater benefits, a conflict of interest arises insofar as these individuals, who prefer a larger to a lesser share of goods, disagree over the right way to divide them.[40] Justice assumes the role of adjudication.

---

[34] See Mills, "Ideal Theory as Ideology," 166.

[35] Rawls, *A Theory of Justice*, 75.

[36] See Rawls, "Justice as Fairness: Political Not Metaphysical," 239.

[37] Rawls, *A Theory of Justice*, 18.

[38] Rawls, *A Theory of Justice*, 19.

[39] Rawls, *A Theory of Justice*, 109.

[40] Rawls, *A Theory of Justice*, 109.

Despite disagreement, the circumstances of justice are not so dire that persons need resort to desperate acts, nor are societies under pressure from famine, war, or democratic crisis. In addition, the individuals who place claims upon one another are susceptible to "defects" since they often fall short in knowledge, thought, and judgement.[41] Rawls writes, "their knowledge is necessarily incomplete, their powers of reasoning, memory, and attention are always limited, and their judgment is likely to be distorted by anxiety, bias, and a preoccupation with their own affairs. Some of these 'defects' spring from moral faults, from selfishness and negligence; but to a large degree, they are simply part of men's natural situation."[42]

After modeling the general circumstances of justice, Rawls constructs the standpoint that best represents an ideal moral relation between individuals. He builds this part of the model under the general but weak condition that a just principle should be one that is not produced by means of coercion. Perhaps one of the greatest shortcomings of previous contract theories is their inability to model this guaranteed state of equality. The contract theories found in Hobbes, Locke, and Rousseau, still allow for relations of domination. By adding a veil of ignorance over the original choosers, Rawls's model nullifies "the effects of specific contingencies" by abstracting away from individual identity.[43]

The final condition of the model concerns the conduct of individuals under the veil of ignorance. When the original choosers arrive at a unanimous conception of justice, they are acting as mutually disinterested, agents expressing their rationality in an original position of equality.[44] Once again, Rawls places weak conditions on these agents. Each individual possesses a particular conception of the good though they do not know the details of that good save for preferring more to less.[45] Their rationality, then, consists simply in having a coherent set of preferences and choosing these preferences according to a plan which will bring about more

[41] Rawls, *A Theory of Justice*, 110.

[42] Rawls, *A Theory of Justice*, 110.

[43] Rawls, *A Theory of Justice*, 118.

[44] Rawls attributes his conception of justice as fairness to Kant's notion of autonomy, which forms a tripartite connection with rational nature and morality From a Kantian point of view, a morally worthy action is directed by reason, which is motivated out of respect for the law (*Groundwork*, 439).

[45] Rawls, *A Theory of Justice*, 123.

rather than fewer goods and in a way that coheres with a sense of justice.[46] The original position (as model) offers a simple and clear justification of the two principles of justice, but consider now a forceful objection to this model. While Rawls views the simplicity of the model as one of its greatest virtues, many disagree with his choice to control for the contingencies of history, personal relations, sentiments, embodiment, and so forth.

The veil of ignorance has proven to be one of Rawls's more controversial components of his theory of justice.[47] It is also what is required, in Rawls's view, to ensure a fair procedure, a procedure that in turn justifies the principles as the just or right principles. Without the veil of ignorance, we have lost the point of the model itself. Before considering these objections in greater detail, I first want to issue a constraint on the kind of criticisms that can be reasonably leveraged here.[48] In practical affairs, humans act on a broad range of motivations, only some of which are easily explainable. Every day we see humans exploit contingencies to their own advantage and we also see them act out of sincere benevolence for others. Rawls does not ignore these facts but rather controls for them. What is more, he is clear that the motivations of persons in the original position are not the same and cannot be the same as the motivations of persons who have a sense of justice and who accept the two principles.[49] The original position may tell us something about the theory of conduct, but it does not follow that we can replicate this in actual situations. Rawls writes, "The conception of the original position is not intended to explain human conduct except insofar as it tries to account for our moral judgements and helps to explain our having a sense of justice."[50] Rawls admits that a person could never actually become an original chooser since we cannot have access to this hypothetical situation in our actual state of affairs. An individual may, however, "simulate" the deliberations of the original position by

[46] Rawls, *A Theory of Justice*, 124.

[47] For one compelling critique of the veil of ignorance along these lines see Seyla Benhabib, *Situating the Self: Gender, Community, and Postmodernism in Contemporary Ethics* (Taylor & Francis Group, 1992).

[48] A second smaller object is that Rawls's model presupposes the principles and then constructs a model that would confirm his solution as the right solution. It is an open question as to whether presupposing an outcome of a model is a problem or is even entirely avoidable. This is not easily resolvable and is moreover beyond the scope of my argument.

[49] Rawls, *A Theory of Justice*, 128.

[50] Rawls, *A Theory of Justice*, 104.

consciously following the constraints placed on the model.[51] The success of this simulation in everyday life will prove difficult since "we will presumably find that our deliberations and judgements are influenced by our special inclinations and attitudes."[52] In Rawls's view, this fault is located in persons who struggle to regulate their practical reasoning rather than in the model itself.[53]

In response to Rawls's idealizations, Onora O'Neill argues that we should stop locating the defect of practical reasoning in humans, if only to preserve an ethical sense of well-being. A better ethical theory, according to O'Neill, maintains the formalism of principles that mark liberal theories of rights, but it does away with justifications that appeal to the "choice procedure of hypothetical idealized agents."[54] O'Neill argues that ethical theories so often fail because they ignore social and historical features that shape human conduct and assume "super-human" capacities of practical reasoning.[55] Her argument also aims to safeguard normative theories against those who want to do away with abstraction altogether.

Rawls's model relies on several abstractions. He assumes weak conditions in his model of conduct, conditions which are derived from widely shared intuitions that are somehow expressive of our moral judgements. The problem for O'Neill is that Rawls idealizes human agents and, consequently, moral judgement. Choosers in the original position are hypothetical agents who have capacities human beings lack. Even this may be a useful way of theorizing if it remained at the descriptive level. The problem, according to O'Neill, is that Rawls's account is not a descriptive explanation insofar as it represents the enhanced capacities as a standard ideal for human conduct. O'Neill writes, "we are to think of idealized agents and their flawless compliance with rational norms as admirable and super-human rather than as irrelevant to human choosing, let alone sub-human."[56]

---

[51] See Rawls, *A Theory of Justice*, 104; 119; 127.

[52] Rawls, *A Theory of Justice*, 127.

[53] Rawls, *A Theory of Justice*, 127.

[54] Onora O'Neill "Abstraction, Idealization, and Ideology in Ethics," in *Moral Philosophy and Contemporary Problems*, ed. J.D.G. Evans (Cambridge University Press, 1987), 56.

[55] O'Neill "Abstraction, Idealization, and Ideology in Ethics," 56.

[56] O'Neill "Abstraction, Idealization, and Ideology in Ethics," 56.

Charles Mills follows O'Neill, claiming that abstraction itself is not the problem. Nor is it in the fact of having normative ideals. His issue lies in how Rawls's model idealizes the social world, at the expense of the real. Mills articulates three senses of the ideal. The first is the ideal-as-normative, which according to Mills is non-controversial insofar as most political philosophy provides prescriptive ideals.[57] The second sense is the ideal-as-model, which I have already noted, is one prevailing method in Rawls's theory. Mills's argument against Rawls concerns a third sense of ideal, which he calls ideal-as-idealized-model.[58]

In ethics, we attempt to capture some phenomenon of the natural or social world, P. One way we do this is by describing P's "crucial aspects (its essential nature) and how it actually works (its basic dynamic)."[59] Mills calls this form of abstraction (Q) an ideal-as-descriptive-model. We can push this model further, by producing an exemplar of what P should be (P'), as both a technical-functionalist kind and a moral kind, which is the ideal-as-idealized model. We regard the P' as an extrapolation or a limit of P *or* we regard Q as being slightly deviant from P'.[60] P' can prove an effective model if it is close enough to Q but if the gap between the two models is too large, beginning with P' renders it an ineffective method for explaining the workings of P.

At first, it may seem peculiar that Mills levies his critiques against Rawls's ideal theory, rather than against strong ideal theory, which openly ignores the actualities of the social world and eschews the practicality of ethics. Mills turns to Rawls, I suggest, because the two share the common ground that political philosophy ought to be a practical endeavor. Both appeal to a set of values and ideals and argue ideals are indispensable for guiding just action. Mills's claim is that Rawls's method offers no guidance since it abstracts away from the realities crucial for ameliorating the injustices it cares about and ignores the historical formations that would make sense of the actual workings of injustice in social institutions. Mills takes his criticism a step further and claims Rawls's method operates under an ideology of the racial contract and so

[57] Mills, "Ideal Theory as Ideology," 166.

[58] Mills, "Ideal Theory as Ideology," 167.

[59] Mills, "Ideal Theory as Ideology," 166.

[60] Mills, "Ideal Theory as Ideology," 167.

obscures rather than illuminates unjust power relations, especially those that cut across race and gender.[61]

Both Mills and O'Neill claim that Rawls is idealizing some phenomena of the social world at the expense of the real, though they disagree about what these phenomena are. Mills claims Rawls presumes an incorrect social ontology and so idealizes social institutions, cognitive spheres, and capacities.[62] O'Neill locates the idealization in what Rawls presumes about human agency and motivation.[63] Both these critiques fail to take Rawls at his word when he claims that his model does not presuppose an ontology or a theory of human motivation. In short, even if this criticism were to incidentally hold against some formulation of Rawls's model, they miss the mark as criticism of the Rawlsian approach to non-ideal theory. A more productive critique, I argue, meets Rawls on his own terms by justifying principles on political and not metaphysical grounds inclusive of theories of human nature. In the following sections, I provide such a critique by way of the political realism of Shklar and Williams.

3. Liberal Political Realism

Political realism typically defines itself in opposition to Rawls and Rawlsian political philosophy. It is therefore unsurprising that, if Shklar and Williams are characterized as realists, their work is also regarded as "anti-Rawlsian." I suggest, however, that more is gained in reading Shklar and Williams as broadly compatible with Rawls. This is especially true of Shklar, who despite her criticisms, is on the whole sympathetic of Rawls's approach to political philosophy within her realist framework.[64]

We can see this most clearly in Shklar's later writings on utopia, such as "The Political Theory of Utopia," and "What is the Use of Utopia?"[65] In the latter, Shklar maps out the tradition

---

[61] Laura Valentini diagnoses this as the "paradox of ideal theory" and claims that it offers the most compelling challenge to Rawls's model. See Laura Valentini, "On the Apparent Paradox of Ideal Theory," *The Journal of Political Philosophy* 17, no. 3 (2008), 332-355.

[62] Mills, "Ideal Theory as Ideology," 168-169. See also Mills, *The Racial Contract* (Cornell University Press, 1997).
[63] O'Neill, "Abstraction, Idealization, and Ideology in Ethics."

[64] For a similar reading to mine, see Katrina Forrester, "Judith Shklar, Bernard Williams and Political Realism," *European Journal of Political Theory* 11, no. 3 (2012), 247-272.

[65] Judith Shklar, "The Political Theory of Utopia," in *Political Thought and Political Thinkers*, ed. Stanley Hoffman (University of Chicago Press, 1998), 161-175; "What is the Use of Utopia?" in *Political Thought and Political Thinkers*, ed. Stanley Hoffman (University of Chicago Press, 1998), 175-192.

of utopian literature—from Plato's *Republic* and Thomas Manne's *Utopia* to Jean-Jacques Rousseau's *The Social Contract* and the 18th- and 19th century utopian socialists Charles Fourier, Robert Owen, and Étienne Cabet. Shklar identifies their shared commitments: a vision of harmonious social life absent group conflict and a belief in "expansive human possibilities."[66] Importantly, utopias do not effect change but depict social harmony so as to measure our defective world against its ideals. Much like in the strong utopians I have discussed, Shklar notes that the point of imaginaries is not to bring about a state of peace and well-being. The utopian effect is "intellectual and indirect, not practical and immediate."[67]

Utopianism, Shklar suggests, belongs to what she calls the "prophetic" model of political theory. Alongside the prophetic, she distinguishes a "normative" model, which "tries to demonstrate the positive potentialities of existing forms of government."[68] The prophetic model offers a transformative vision of politics but cannot tell us how to express that utopian vision. Contemporary normative models, by contrast, are essentially reformist insofar as they supply formal and critical procedures for constitutional democracies, procedures already embedded in our lives but simply not realized. Rawls's *A Theory of Justice,* she notes, is the clearest contemporary illustration of the normative model.

Like Shklar, Williams too shares a commitment to a normative model of political philosophy that vindicates liberal values. My aim in drawing out these similarities is not to recast Shklar and Williams as Rawlsians. Their differences, as I will demonstrate, are equally significant. However, Shklar and Williams offer the most compelling liberal reformulations of Rawls, reformulations that remain broadly faithful to, or at any rate appreciative of, the Rawlsian project. In particular, their realist visions preserve liberalism and through it seeks to justify values that are already present in fragmentary form but not yet fully realized.

The difference, then, is that Shklar and Williams are primarily skeptical of a normative theory that emerges solely from moral reason. We may initially fear the uncertainty of an ethical life untethered to a universally valid and rational order, but our fear would be unfounded since this uncertainty has always been a part of our political condition. Ethical life is constructed by

[66] Shklar, "What is the Use of Utopia," 177.

[67] Shklar, "What is the Use of Utopia," 178.

[68] Shklar, "What is the Use of Utopia," 187.

more than rational argument, which only takes us so far, since there is more to our social states than coherency and cogency. To insist on the limits of philosophy is not to slight it, but to demand that it understand its situation more truthfully.

One of the greatest misconceptions about political realism concerns not simply what it doubts but *how* it doubts. When the political realist is faced with an evaluative or normative concept, she grasps that reality is elusive and that hidden ignorance can inspire false intellectual self-assurance.[69] She holds an unconventional view of accepted social beliefs and judges with an attitude of doubt. The fanatic, the dogmatist, the uncompromising moralist, always finds answers in his own political truth, without which, he warns, we devolve into a world of doom, disorder, and unrestrained inclinations. Political skepticism on the other hand, presents a compromising sort of doubt by suspending judgement in the face of intractable disagreement. Now, of course, skepticism can become its own kind of dogmatism. As Shklar notes, a misanthropic skeptic often perpetuates horrific cruelties and a society of oppressive skeptics is easy to imagine if it is not careful to restrain itself.[70] To overcome this, Shklar argues that skepticism must be inclined towards toleration insofar as it does not rest permanently on any single belief.[71] At the same time, if one's skepticism results in taking a step back from conflict, it can easily give way to quietism or nihilism. Looking to Montaigne's skepticism, Shklar notes he expresses "a refusal to accept either the comforts of political passivity or of Machiavelli's platitudes. [Montaigne] did choose sides in the civil war, did his duty, and kept a sharp eye on the mayor."[72]

Shklar also sets limits on normative theory, whose task, she writes, has always been

> The elucidation of common experience, the expression of what is inaccurately known to groups of people at any time. It is the re-examination of inherited ideas, their adaption, and even their utter rejection. It is not a work of discovery but one of illumination through discussion. We all carry with us a mixed bag of *idées reçus,* and in order to travel with it through an everchanging world we must shift it around occasionally—drop something here and add something there. Above all, we share a variety of common

[69] Judith Shklar, *The Faces of Injustice* (Yale University Press, 1990), 20.

[70] See Judith Shklar, *Ordinary Vices* and Shklar, "The Liberalism of Fear," in *Liberalism and the Moral Life*, ed. Nancy L. Rosenblum (Harvard University Press: 1989), 21-38.

[71] Shklar, "The Liberalism of Fear," 25.

[72] Shklar, *Ordinary Vices,* 34.

> encounters and reactions, and to give these a coherent form, a clear voice, is the task of political theory.[73]

Shklar's description of political theory as a practice of reexamination through discussion resonates with Williams's claim that, in ethics, "reflection can destroy knowledge."[74] In the case of ethical reflection, the concepts that once guided us through the social world no longer perform their function.[75] Skepticism therefore holds that "ethical knowledge is not necessarily the best ethical state."[76] Seeing as reflection would likely reveal the radical contingency of our ethical representations, some may conclude that it risks effacing itself before it can even leave the ground. Reflection too early on may throw us into a state of what Amia Srinivasan calls "genealogical anxiety."[77] Once we comprehend the contingent origins of what we reflect on, we grow fearful that the authority of our ethical representations will cease its grip, that it will somehow cast doubt on the legitimacy of our ethical concepts.

Williams assures us that even when ethical reflection engenders doubt, it at once stirs up confidence.[78] We say we are confident precisely because we cannot know all the changes that will affect us in the future. Confidence sometimes looks like hope, though, it may also look like the rueful admission that people and institutions will behave badly. We may, for example, be confident in the persistence of corrupt and hypocritical politicians, and equally confident in an ethical undertaking that tries to restrain cruel or violent ones. This is the basic idea behind Shklar's barebones liberalism.[79] Practical confidence is, in some sense, a practice of truthfulness.

---

[73] Judith Shklar, *Legalism: Laws, Morals, and Political Trials* (Harvard University Press, 1964), 28.

[74] Bernard Williams, *Ethics and the Limits of Philosophy* (Harvard University Press, 1985), 148.

[75] Williams, *Ethics and the Limits of Philosophy*, 167.

[76] Williams, *Ethics and the Limits of Philosophy*, 168.

[77] Amia Srinivasan, "Genealogy, Epistemology, and Worldmaking," *Proceedings of the Aristotelian Society* 119, no. 3 (2019), 127-156.

[78] Williams, *Ethics and the Limits of Philosophy*, 171. For a more detailed discussion of confidence in Williams, see Edward Hall, "Contingency, Confidence, and Liberalism in the Political Thought of Bernard Williams" *Social Theory and Practice* 40, no. 4 (2014), 545-569; Hall, *Value, Conflict, and Order: Berlin, Hampshire, Williams, and the Realist Revival in Political Theory* (Chicago University Press, 2020).

[79] Shklar first introduces this phrase in *Legalism* and returns to it in her seminal essay "The Liberalism of Fear."

As Williams writes, it stems from a pessimism of strength rather than "the weakness of self-deception and dogmatism."[80]

One risk with confidence is that it can end up sounding like a virtuous disposition we take up in private or with an interlocutor. This is only partly true. Williams notes that confidence is basically a social phenomenon—it exists within a society and requires social confirmation.[81] Still, some are unconvinced confidence can get us very far. John Cottingham describes Williams's project as forcing a resolution on a "fundamentally irresolvable tension between ethical conviction and contingency."[82] Realists admit that part of ethical conviction stems from being convinced in a way that goes beyond rational argument, and so, in some sense passes over us in a way that makes our actions feel inescapable.[83] Sometimes those actions appear to us contingently, and in such a manner that we cannot speak to their obscure causes and effects.[84] Equally so, ethical conviction can be strengthened upon reflection. To echo Shklar, we (as a society) choose to drop certain convictions and pick others up. A politically realist skepticism stirs up doubt through reflection while also confidently affirming some ethical modes of living over others. This practice of illumination is never complete since confidence does not amount to objective knowledge. A skeptic does not irrevocably seal in values. She doubts to keep the ethical conversation in motion.

A naïve reader of Williams's realism may misconstrue his critique of political moralism as a critique of normative ideals writ large.[85] Such a reading would be incongruous with the deeply ethical dimension of his broader philosophical work and Williams's belief that almost everyone

---

[80] Williams, *Ethics and the Limits of Philosophy*, 171.

[81] Williams, *Ethics and the Limits of Philosophy*, 170.

[82] John Cottingham, "The Good Life and the Radical Contingency of the Ethical," in *Reading Bernard Williams*, ed. D. Callcut (Routledge Press, 2008), 25-43; see also Mark Jenkins, *Bernard Williams* (Routledge Press, 2006).
[83] Williams, *Ethics and the Limits of Philosophy*, 169.

[84] Bernard Williams, *In the Beginning was the Deed* (Princeton University Press, 2005), 13.

[85] See David Estlund, "Anti-Anti-Moralism," in *Utopophobia,* 40-61. See also political realists who overlook the ethical dimension of Williams's work, such as William Galston, "Realism in Political Theory,"; Enzo Rossi and Matt Sleat, "Realism in Normative Political Theory," *Philosophy Compass* 9, no. 10 (2014), 689-701; Sleat, "Bernard Williams and the Possibility of a Realist Political Theory," *European Journal of Political Theory* 9, no. 4 (2010), 485-503; Matt Sleat, "Legitimacy in Realist Thought," *Political Theory* 42, no. 3 (2014), 314-337. For example, Rossi and Sleat write, "realists posit a dichotomy between the realm of human action that is appropriately regulated by morality (or ethics—we use those terms interchangeably for now) and the realm of politics, which requires separate norms" (691). For Williams, however, these terms are not interchangeable. Ethics exists as a larger category with morality as one of its subsets.

entertains ethical considerations in daily life.[86] The perils of morality only arise when individuals within the morality system mistake it to be the absolute system, that is, the only one that ultimately matters, even to those who do not accept its rules.

Williams's ethical outlook pushes against the morality system's tendency to reduce everything to obligations and then to devalue all other considerations in political deliberation. One way moralism gets around the uncompromising nature of its deliberative demands is by turning all considerations into obligations. The maxim Williams assigns to this aspect of moralism is "*only an obligation can beat an obligation*."[87] Suppose I am under an obligation to obey a law even if I believe it is unjust. From the point of view of morality, I can only permit myself, morally speaking, to break my obedience to the law under the pretense of some other obligation, such as my obligation to the victims of this unjust law. What these competing, particular obligations require then, is clarification with respect to some general obligation, such as the obligation to uphold the stability of institutions, or the obligation to promote human flourishing. Williams refers to this as the "obligation-in, obligation-out" principle. Depending on which general obligation takes highest priority, I am either bound to or recused from my obedience to the unjust law. This peculiar principle determines how we assign blame to ourselves and to others. So long as I first ask myself whether I am under an obligation and then follow that obligation correctly for any individual case, I will not stand accused of blame. I can assure myself that, even if I have done harm, I have not acted wrongly. Here, Williams notes one significant difference between moralism and a broader sense of ethical life. When we assign blame, "the thought *I did it* has no special significance; what is significant is whether I voluntarily did what I ought to have done."[88] In the face of a difficult choice, my conscience escapes unscathed. If I am fully captured by the morality system, I may feel pressure to keep my conscience pure, that is, moralism may place me under a general obligation to not waste time on morally indifferent actions.[89] For example, I may feel immense guilt for tending to my garden (literally or metaphorically) when I could be

[86] Williams, *Ethics and the Limits of Philosophy*, 174.

[87] Williams, *Ethics and the Limits of Philosophy*, 180.

[88] Williams, *Ethics and the Limits of Philosophy,* 177.

[89] Williams, *Ethics and the Limits of Philosophy*, 181.

volunteering at the local food pantry. Even so, the trouble is that moralism's tendency to reduce all action to moral action easily comes to "dominate life altogether."[90]

Williams notes that political moralism makes a claim on us in virtue of our rational nature and therefore sees itself as more demanding than the law of any state. Williams's critique of moralism on this point is anticipated by Shklar's critique of analytical positivism and natural law in *Legalism*. She here argues that so long as rules reign, "the structure of legal and moral reasoning is identical."[91] We can therefore understand moralism within a broader political environment of rule-oriented thinking where normative action is understood in terms of following rules and precepts and where laws and morals share a vocabulary of obligation. Though the disagreements between positivists and natural law theorists run deep, Shklar argues that both ultimately treat law as a distinct and self-evident institution immune from politics and manned by a judiciary that attempts to maintain fairness and impartiality.[92] Both theories lay down the legal conditions under which power can be justly exercised and treats politics as nothing more than applied legality. Shklar calls this view legalism, which she defines as "the ethical attitude that holds moral conduct to be a matter of rule following, and moral relationships to consist of duties and rights determined by rules."[93]

Moralism and legalism are most united by the idea of justice. Even the leading proponent of legal positivism, H.L.A. Hart, recognizes the resonance, writing "justice is both a virtue especially appropriate to law and the most legal of the virtues."[94] Shklar characterizes justice as a legalistic ethic insofar as it is expressed in the thought and conduct of individuals as a loyalty to rules, and an acceptance of obligation "under a system of principles."[95] As a policy of legalism, justice requires a general rule to be applied fairly and impartially to an individual instance.[96] A person is just when she acts in accordance with an agreed upon system of rules rather than

---

[90] Williams, *Ethics and the Limits of Philosophy*, 182.

[91] Shklar, *Legalism*, 61.

[92] Shklar, *Legalism*, 35.

[93] Shklar, *Legalism*, 1.

[94] H.L.A. Hart, *The Concept of Law*, (Oxford University Press, 1961), 7.

[95] Shklar, *Legalism*, 113.

[96] Shklar, *Legalism,* 114.

abiding by a specified rule or act. This "habit of mind," as Shklar puts it, is best understood as a virtuous attitude of agreeing to a set of rules and following those rules, whatever they may be.[97] To curb one's set of practical identities, to adopt a disposition to fairness and impartiality, to give an individual her due, is for Shklar, an expression of one's character.[98] Under the policy of legalism, justice arises when individuals place claims on one another and are adjudicated by distributing power among individuals fairly and impartially according to a system of rules.[99] A legal institution is deemed just when a judge expresses a set of principles or follows a set of rules fairly and impartially. She ensures that every person is given her due.

Unlike Rawls, who believes justice can bring about a more egalitarian society, Shklar argues that the excessive formalism of legalistic rule-following struggles to maintain a stable institutional grip over a pluralistic society. Shklar writes:

> Even the most impartial and fair-minded people, and especially those who are in a position to impose political or judicial decisions among us, will seem unjust sooner or later…That is why just behavior does not necessarily lead to results recognized as just by those affected. There is bound to be too much disagreement to make justice perfect in practice, too many mutually exclusive standards to make justice either a simple or clear policy.[100]

There is an ambiguity in justice due to an asymmetry between our expectations of how cases should be adjudicated and how they actually are. We react acutely to injury, but this does not mean an injustice has occurred. Those who accept the given legality system will in turn accept the outcome as just and will see the conflict as resolved. For those outside the system, in competing systems, or who interpret the facts differently, injustices will proliferate. Shklar therefore skeptically asks whether justice is really capable of accommodating the diversity of a polity into its general rules of conduct. Though rules are not typically arbitrarily created, we nevertheless construct them. When legalism essentializes distinctions between justice and injustice, it fails to take stock of the changing social world.

---

[97] Shklar, *Legalism*, 114.

[98] Shklar, *Legalism*, 115.

[99] Shklar, *Legalism*, 117.

[100] Shklar, *Legalism*, 117.

Apt as her critiques may be, Shklar's occasional polemics overlook a potential ally in Hart who embraces the indeterminacy of rules through his notion of "the open texture of law."[101] He writes: "We should not cherish, even as an ideal, the conception of a rule so detailed the question whether it applied or not to a particular case was always settled in advance."[102] A world of "mechanical jurisprudence" may be an easy world for it would shelter us from the uncertainty of choice and loss in the hard cases, and yet, Hart remarks, "this world is not our world."[103] He agrees then that there is a limit to how well generalized or abstract language can guide us in social life. At the same time, he criticizes extreme rule-skepticism for characterizing rules as myths and for reducing the law to mere decisions of the courts or predictions on those decisions.[104]

I agree with Hart that rule-skepticism often conflates our inability to ensure determinate outcomes with a failure of rules in their entirety, and in so doing, becomes just as wedded to an unattainable ideal as in a formalist's utopia. Shklar takes heed to restrain her skepticism before it becomes its own kind of formalism and accepts that Hart's skeptical view of politics is congenial with her own.[105] She does not deny the existence of rules nor their ability to meaningfully determine outcomes. Her skepticism aims to demarcate legalism's limits of use, to show that justice can only take us so far or resolve so much. For example, demands for "social justice" or "political justice" are less complicated and more accommodating when we call them simply what they are: equality and freedom.[106]

In a much later book, *The Faces of Injustice,* Shklar notes that, historically speaking, forcing social resolutions to conform to a general system of rules "tends to backfire."[107] We may believe that they reflect our intuitions and reduce our cruelties, but as Shklar sees it, they are more likely

[101] Hart, *The Concept of Law*, 130.

[102] Hart, *The Concept of Law*, 128.

[103] Hart, *The Concept of Law*, 128.

[104] Hart, *The Concept of Law*, 136.

[105] Shklar, *Legalism*, 7.

[106] Shklar, *Legalism*, 119.

[107] Shklar, *The Faces of Injustice*, 26.

to "redirect and formalize our ferocity."[108] She attributes this to the uncompromising nature of legalism, and its attempt to, in some sense, adhere to the general obligation of applying a rule to an individual case. Instead, she advocates for a policy of "semi-justice," in which conflicting goals, such as a need for justice and a need for forgiveness, can exist in compromise.[109]

4. Putting Cruelty First

Rawls's original position sets out the moral conditions of co-existence and specifies the kinds of principles, laws, and institutions that free and equal people would reasonably agree to. On this view, a society in which power is rightfully exercised is one and the same as a society that satisfies the two principles of justice. Political realists, by contrast, argue that we ought to analytically distinguish between the rightful exercise of power and liberal aspirations to social justice, such that a state can conceivably secure the former without realizing the latter. Williams refers to this more minimal form of justification as the "basic legitimation demand," which is basic in the sense that it seeks to secure the conditions of legitimate power by establishing order, safety, and trust against a backdrop of war, terror, and domination. A basic justification takes a second-person form insofar as it offers reasons to each subject for how a given exercise of power is meant to alleviate or replace the fear of coercion, suffering, humiliation, and death that would arise in its absence.[110] A justification must therefore explain how power functions as a solution to the problem of fear without relying solely on coercion itself.

As a minimal ethical condition, realism adopts the maxim that might does not imply right.[111] This maxim, however, speaks not only to ethical concerns but also to prudential ones, since a state that secures agreement through power alone will not secure stability. Of course, order and security could prevail under coercive rule, but coercion is nevertheless ill-suited to generating relations of trust. This is because trust concerns the normative expectations we place on others that they will do what they are supposed to do. Although we typically understand trust in ethical and interpersonal terms, such as when I trust my friend to keep my secret, it also encompasses

[108] Shklar, *The Faces of Injustice*, 26.

[109] Shklar, *The Faces of Injustice*, 122.

[110] See Stephen Darwall, *The Second Person Standpoint: Morality, Respect, and Accountability* (Harvard University Press, 2009).

[111] Williams, "Realism and Moralism," 4.

default and prudential expectations, such as when I trust the government to keep my information safe.[112] These background expectations enable individuals to conduct their daily lives with a degree of assurance that things will go according to plan and that others will act as anticipated. Without trust's assurance, social life becomes characterized by uncertainty and the pervasive fear that accompanies it. This is not simply a theoretical claim but can be backed up by plenty of empirical examples.

One contemporary way to understand legitimate power is through human rights discourse. Within this framework, if a state violates a human right, it means it has wronged an individual or group in a way that approximates war or severe domination. This is why, for Williams, human rights are not aspirational but matters of urgency.[113] To accuse a government of violating human rights must be met with the utmost sense of responsibility since it is the gravest political demand one can place on the state, and is one that immediately calls others to action.[114] As such, Williams excludes certain positive rights, such as the right to work or the right to abortion and focuses instead on the most grievous injuries such as torture, surveillance, arbitrary arrest, murder, and execution.[115] Williams does not therefore claim that positive rights hold little importance. They demand political urgency and evoke distinct normative responses. Unlike other rights violations, human rights violations are politically unjustifiable since they *ipso facto* fail the critical theory test. States that have the resources or political coordination to secure loftier political ambitions or that adhere to different conceptions of human flourishing, may raise the standards of what counts as a right within their own jurisdictions.

Williams's universal conception of human rights violations rests on a morally austere conception of human rights. Philosophical and political problems emerge in hard cases, when a thin line separates brute coercive power and legitimate authority.[116] For instance, liberals would not accept as legitimate a theocratic regime that subordinates women. The question is whether

---

[112] For her discussion of trust, see Margaret Urban Walker, *Moral Repair: Reconstructing Moral Relations after Wrongdoing* (Cambridge University Press, 2006), 72-109.

[113] Bernard Williams, "Human Rights and Relativism" in *In the Beginning Was the Deed*, ed. Geoffrey Hawthorn (Princeton University Press, 2005), 64.

[114] Williams, "Human Rights and Relativism," 72.

[115] Williams, "Human Rights and Relativism," 62.

[116] Williams, "Human Rights and Relativism," 63.

we should consider those practices human rights violations.[117] Part of the solution may be located in what Williams calls the "critical theory test," where the subjects under power are asked to reflect on how they came to accept the justification in the first place and to consider, in light of that reflection, whether they still regard the power as legitimate.[118] More pragmatically, when we cannot promote such consciousness-raising—either ethically or feasibly—the best course of action may be to denounce the state's practices while limiting intervention. As Williams warns, "If the case is one which looks less like unmediated coercion, the victims may not think they are victims, and the intervention may be difficult to distinguish from ideological imperialism."[119] Williams's historical sense allows him to recognize that an outside state's intervention for the purposes of offering reprieve from human rights violations can easily become the problem by producing new forms of human rights violations. Part of the response to perceived violations will thus involve considerations about what actually can be done rather than what ought to be done about those practices. This will involve, as I have been arguing, not simply ethical or moral considerations but prudential ones, among many others.

In Williams's view, it is possible to imagine a range of legitimate states that are non-liberal. At the same time, he maintains that the coherence of a legitimate non-liberal state "does not exclude the possibility that there might be circumstances in which the only way to be [legitimate] involved being liberal."[120] This suggests that, although non-liberal justifications of political power are intelligible in theory, historical and social contexts make it such that the only legitimate state is one that secures liberal values, such as freedom and equality. Williams reiterates this point writing, "Now and around here, together with historical conditions permit

[117] Williams, "Human Rights and Relativism," 71.

[118] In *Truth and Truthfulness* Williams provides a far more substantive critical theory test, though for now it suffices to view the test as a process of reflection for some group for the purposes of distinguishing between authority and power. If, upon reflection, what was once accepted as a justification is now understood to be a product of coercive power, the authority it once had over the group becomes de-stabilized, and not simply insofar it is thrown into doubt, but it now is seen as unjust, illegitimate, and no longer accepted as a source of authority. When a state is in want of some justifying explanation of its authority Williams contends there is nothing to be said to those groups to explain why they should not revolt . See Williams, "Truthfulness, Liberalism, and Critique," in *Truth & Truthfulness: An Essay in Genealogy* (Princeton University Press, 2002), 206-232.

[119] Williams, "Human Rights and Relativism," 73.

[120] Williams, "Realism and Moralism," 4.

only a liberal solution: other forms of answer are unacceptable."[121] The apparent tension between these claims might make us wonder how legitimacy can be both compatible with non-liberal forms of rule and, in certain circumstances, require a liberal justification.

The tension here is easily resolved by recognize that what counts as an answer to the basic legitimation demand influenced by history. From our present standpoint, it is all too easy to import more robust liberal values in the basic legitimation demand, a mistake committed by much of Rawlsian political philosophy. A historical perspective, however, reminds us that liberalism is a comparatively recent arrangement on the political *mise en scène* and one whose success in securing its basic values of freedom and equality has been repeatedly contested, not least under the pressures of fascist authoritarianism. We must take care to ensure that the structural conditions we assume in the basic legitimation demand are indeed minimal conditions for the possibility of politics as such. At the same time, a historical sense explains why, despite the existence of intelligible, non-liberal legitimations in the past, and despite its possibility in the future, liberalism is the only permissible answer to the basic legitimation demand "now and around here."[122] In this respect, we can describe the emergence of liberalism in the late seventeenth century as "intellectually irreversible," not because its values are immune to erosion or loss, but because it has extended the standards of disadvantage, reasonable fear, assaults on the self, and unacceptable exercises of power.[123]

Does such a formulation of liberalism lead us to reinterpret Rawls's principles of justice as a historically situated answer to the basic legitimation demand, such that only something approximating justice as fairness would suffice for political legitimacy? I suggest that it does not. Although liberalism extends the standards of the basic legitimation demand, there are varieties of liberalism, ranging from less to more idealized. The task, then, is not to identify the most demanding liberal theory, but to articulate a version of liberalism suited to the kinds of people we

[121] Williams, "Realism and Moralism," 8.

[122] Williams, "Realism and Moralism," 8.

[123] My interpretation of liberalism as an extension of the basic legitimation demand is supported not only by Williams's explicit remarks in his political writings, but also by his account of one distinctive use of history in *Truth and Truthfulness*. There, Williams suggests that historical developments can extend the structural needs identified in his State of Nature story and illustrates the point through the emergence of historical time and historical truth. Indeed, I will discuss this point in greater depth in Chapter Four, though for now it is helpful to show how Williams's political philosophy aligns closely with his broader genealogical approach. Indeed, we can read "Realism and Moralism" as a proto-genealogy of liberalism itself.

are and to the political aims that can reasonably be pursued in light of our history. For this account, I argue we ought to turn to Shklar's "liberalism of fear".

In his review of *Ordinary Vices,* Williams describes Shklar as navigating a viewpoint that affirms the modern liberal state as, on the one hand, improving its citizens and laws, and on the other, operating as an impersonal system of law and administration.[124] Rather than seeking to reconcile the two conceptions, *Ordinary Vices* serves only to suspend the reader between a collective ethical life worth living and an impersonal regulatory framework that permits individuals to get on with their lives with limited interference from the state and others. Five years later, with the publication of Shklar's essay "The Liberalism of Fear" there is a discernable shift in the aims of her political theory. It is now more pessimistic—or at any rate, less hopeful—than her earlier work, though paradoxically, more determined to construct a positive framework for political thought. What began in *Ordinary Vices* as a general intuition to "put cruelty first" evolves in "The Liberalism of Fear" into a first principle for liberal political philosophy: putting cruelty first. For Shklar, when someone lives in fear of psychological or physical cruelty, she cannot in any meaningful way develop her own interests. Instead, her actions are by the need to avoid cruelty, thus placing her under the control of someone or something else. In this way, her choices are guided by the intentions of others even when those intentions conflict with her own.

The liberalism of fear aims for what Shklar calls negative egalitarianism, "an obvious corollary of putting cruelty first," which is rooted in a suspicion of the effects of inequality and the unfreedoms such inequalities produce. And because the cruelty principle concerns the conditions of possibility for cruelty, it blends elements of both negative and positive freedom. Laws, institutions, and the demand for justice, can quicken the effects of cruelty. Here Shklar is thinking of the excesses of governmental power expressed in official bureaucratic agents, military, paramilitary, and police.[125] Equally so, laws and institutions can make collective life better, insofar as they ensure a dispersion of power that mitigates cruelty and so enables the conditions of freedom.

Shklar's liberalism of fear thus extends the standards of the basic legitimation demand. A just exercise of power must now not simply secure the conditions of order, safety, and trust, but

[124] Bernard Williams, "Resisting the Avalanche," *London Review of Books* 7, no. 10, 1985.

[125] Shklar, "The Liberalism of Fear," 29.

also the political conditions necessary for the exercise of personal freedom, understood as the ability of adults to make effective decisions about their lives without fear, and consistent with the like freedom of others.[126] In this way, the liberalism of fear places comparatively demanding requirements on political legitimacy. It charges the state not merely with maintaining order, but with securing two distinctively liberal conditions that structure how power may be justly exercised: freedom and equality. These requirements exclude bare or minimal solutions to the basic legitimation demand and render unacceptable certain social arrangements including deeply entrenched hierarchies, insofar as they subject individuals to unmediated forms of coercion incompatible with freedom from fear.

The liberalism of fear, then, elevates the avoidance of cruelty and fear to the status of a basic norm within political life by aiming to construct a government "for men as they are, not as they might be," and so responds to the basic intuition that most people would avoid cruelty, and the fear cruelty produces.[127] A "liberal government for bad characters" does not promise freedom will improve the liberal person, but that it will remove the worst impediments to ethical life.[128] For some, the principle's austerity is nothing short of dystopian and threatens confidence in our liberal convictions. In response, I suggest we distinguish between a morally austere principle and one that is politically ambitious. Although Shklar's justificatory model refrains from supplying a moral principle it is still politically ambitious and should not diminish confidence in the liberal project itself.

Even more than avoiding moral consensus, the cruelty principle deliberately avoids foundationalism of any kind, including ethical naturalism, with which it might otherwise be confused. The fear it mitigates is not one stemming from the fact of being alive, but a fear produced by a political relation between individuals and governmental power, and by "arbitrary, unexpected, unnecessary, and unlicensed" state-sanctioned public cruelty.[129] Since institutions typically support and sustain abuses of power, Shklar's principle is directed at the governmental agents who abuse their institutional power: official bureaucrats, police, military, paramilitary, and

[126] Shklar, "The Liberalism of Fear," 21.

[127] Shklar, "The Liberalism of Fear," 29.

[128] Shklar, *Ordinary Vices*, 236.

[129] Shklar, "The Liberalism of Fear," 29.

even private corporate actors. For Shklar, given the "inevitability of an inequality of military, police, and persuasive power which is called government," abuses will occur, and will likely disproportionally fall on the "permanent minorities."[130] Grounded in the lessons of political history, Shklar rightly assumes state actors will frequently enact lawless and brutal behavior absent restraint on their actions.[131]

Despite her skepticism of governmental power, Shklar claims her cruelty principle is expressed through institutions vis-à-vis the legal protection of equal rights.[132] On Shklar's account, institutions are pragmatic tools tasked with enabling individuals with rights. "The first right," Shklar claims, "is to be protected against the fear of cruelty. People have rights as a shield against this greatest of public vices."[133] Shklar is wise enough to recognize that preventing the overreach of governmental or corporate power does not necessarily require smaller government or fewer laws. Instead, it calls for laws and institutions that effectively diffuse power, support citizens in conducting their government, and so relieve "the burden of fear."[134]

The integration of cruelty into a liberalism of rights is no less than an intellectual advancement worth vindicating and one we should not take for granted. As Williams notes, a strength of the liberalism of fear is that it reminds people "of what they have got and how it might go away."[135] What is more, in pursuing idealized notions of equality and freedom, we might overlook that, even within constitutional democracies, there are those who lack basic protections from fear and cruelty and for whom ambitious ideals hold no meaning. By failing to attend to the most politically vulnerable we risk effacing the legitimacy of a constitutional democracy and may lose liberalism entirely.

---

[130] Shklar, "The Liberalism of Fear," 27.

[131] Shklar, "The Liberalism of Fear," 27-28. In *Political Liberalism*, Rawls does acknowledge the coercion underlying political power since government institutions and laws are generally backed by the legitimate threat and implementation of sanctions. At the same time, he writes that in a constitutional regime, "political power is ultimately the power of the public, that is, the power of free and equal citizens as a collective body." See John Rawls, *Political Liberalism* (Columbia University Press, 1995), 135.

[132] Shklar, "The Liberalism of Fear," 37.

[133] Shklar, *Ordinary Vices*, 237.

[134] Shklar, *Ordinary Vices*, 238.

[135] Bernard Williams, "The Liberalism of Fear," in *In the Beginning Was the Deed*, ed. Geoffrey Hawthorn (Princeton University Press, 2005), 60.

5. Thickening Cruelty

Although cruelty addresses the need for a shared foundation for moral intuitions in a modern society, one critique is that its minimal ethic may be too austere to guide our action and may equip us with too few ideas to meet our needs. Moreover, we are unlikely to agree that cruelty ought to be our first political principle, since realistically, people think in a broad range of needs and interests, some of which will not coincide with putting cruelty first. It is in part for this reason that Shklar stresses the need to move beyond mere intuitions and towards a justification of the cruelty principle. For a compelling justification, it must make sense to those it seeks to persuade. This entails that it will be realizable insofar as it is capable of guiding political action.

If our aim is to construct action-guiding political principles, then we need ethical concepts that are both shaped by the world (insofar as they are practice-dependent), and ones that are world-guiding (insofar as they provide reasons for action), which Williams refers to as "thick ethical concepts."[136] According to Williams, our evaluations extend beyond thin ethical concepts, such as right or wrong, and towards thicker concepts, such as sincerity, care, courage, selfishness, imprudence, or cruelty. What distinguishes a thin from a thick concept is that the latter expresses a union of fact and value such that we draw on these terms not only to suggest a valuation of a person or action, but also to lend some descriptive weight to that evaluation.[137]

Calling someone's actions cruel evaluates the action as wrong but also specifies a particular way of acting wrongly. Cruel acts evoke something more specific, such as an agent being indifferent to the suffering of others even when he could step in to help. In addition to providing a reason for action, the concept also conjures up an image of a particular action that would correctly apply to the concept and situation. If cruelty is negatively evaluated and I can image specific cruel actions, then I should (and presumably would want to) either avoid those actions or

---

[136] Williams, *Ethics and the Limits of Philosophy*, 129-130. Williams describes thick ethical concepts that contain both negative and positive evaluations, though he also notes the presence of *counter-ethical* concepts, which he claims run counter to and are parasitic on the ethical. It is noteworthy that Williams chooses cruelty as the paradigm of the counter-ethical. Following Nietzsche, Williams claims cruelty is parasitic on the sensibility of sympathy (*ELP*, 13-14).

[137] Williams goes so far as to say that a thick ethical concept cannot make sense, and so cannot be judged or evaluated, if both its descriptive and evaluative contents are not grasped. For example, an outsider may judge a correct or incorrect application of a concept with considerable regularity, though she may be prone to misapply the concept in certain instances since she will fail to grasp its evaluative dimension. Thick ethical concepts are thus grasped from both an internal and external point of view. For more on this see, Williams, "Internal and External Reasons" in *Moral Luck* (Cambridge University Press, 1982); and, Hart, *The Concept of Law*, 89-91.

do my best to mitigate them. Now, as Williams suggests, a reason for action does not guarantee the action itself. There may be other reasons that cause me to act cruelly, and these reasons may not be good reasons or may be supplanted by more rudimentary emotions, such as fear. A negative or positive evaluation of a concept, paired with a clear image of what that looks like in practice, will operate as the force of reason that either propels a thick ethical concept into action or keeps it at bay. In the right circumstances, with the right balance of considerations, and for the right agents, thick ethical concepts will indeed guide action.

A description of a thick ethical concept does not guarantee that individuals will easily or unanimously agree on its correct application. In fact, the distinctive element of thick ethical concepts is that both their evaluative and descriptive meanings rest on localized information. What is cruel in one historical configuration or geographical setting might be unreflectively accepted as a fair practice in another. In many cases, agreement on thick ethical concepts will seem intractable. Such a contestation of a concept's use even of the most intractable kind does not mean it is no longer controlled by the facts. Disagreement, as Williams suggests, reveals the opposite: the world is guiding us towards one meaning instead of another and that meaning is worth arguing over.

Part of why it is difficult to characterize the cruelty principle in abstract terms is because the kind of action it guides will depend on those thicker ethical concepts operative at any given historical juncture, alongside other political considerations. This is all to say that the cruelty principle does not provide a clear prescription as to what to do next. Its content is minimal, its form is austere. It does, however, give us a general direction towards which to direct our political aims, and calls on us to use our own judgement. Part of this will require making cruelty a thick ethical concept, or better yet, a thick counter-ethical concept. It will also involve cultivating and curtailing other counter-ethical concepts. For example, hypocrisy, which runs counter to authenticity and sincerity, engenders a form of dishonesty that makes cruelty easier.[138] At the same time, we may do well to ignore hypocrisy, or at least to understand that its occurrence often stems from confused intentions or conflicts of loyalty.[139] By becoming less outraged by hypocrisy we will resist the cruel responses it invites.

---

[138] Shklar, *Ordinary Vices*, 12.

[139] For her discussion of hypocrisy, see Hannah Arendt, "The Social Question," in *On Revolution* (Penguin Books, 1963).

Keeping one's eye out for cruelty can make a person misanthropic, a counter-ethical concept that is parasitic on benevolence. The danger with misanthropy is that the anger it creates too often justifies violence against humanity.[140] And yet, an honest approach to how humans behave can motivate us to put up legal and political safeguards that would attenuate the effects of cruelty.[141] Once again, Shklar acknowledges that it may not be feasible or desirable to altogether rid ourselves of the vice. Like Montaigne, she turns to friendship so as to resist the "avalanche of disgust which can at any moment overwhelm anyone."[142]

In particularly dystopian societies, cruelty will not generate a strong negative evaluation at all. In most societies, considerations of cruelty will conflict with other considerations. In *Ordinary Vices,* Shklar focuses on Nadine Gordimer's *Burger's Daughter*, which is set in apartheid South Africa. In a striking scene, Rosa, the protagonist, is confronted with a Black man beating a battered donkey, but she cannot bring herself to stop him or do anything about it and so flees. For Shklar, Rosa is torn between calling the police and putting cruelty first or turning away and putting political oppression first. In this, Shklar writes, "we can see her strength of character especially in her refusal to pretend that either decision would have been obvious or flawless. Putting cruelty first might help one to decide who the victim at any moment is, but not without some very real doubts and uncertainties."[143] Having standards, Shklar reminds us, will not save us from doubt or ethical anguish when an action must be taken, but it just might help us get on in political life and to do as little harm as we feasibly can.

One reason why ideal theory derides politics is due to a categorical confusion between contingency and arbitrariness.[144] As I have mentioned, there is a concern under Rawls's model that once we understand our convictions as products of historical conditions, we will fail to secure ethical agreement. For the realist, contingency does not give us license to abrogate our convictions. It simply requires justification from within political life for power to make sense. As Williams writes, "We may be able to show how a given practice hangs together with other

[140] Shklar *Ordinary Vices,* 192.

[141] Shklar, *Ordinary Vices*, 193.

[142] Shklar *Ordinary Vices*, 215.

[143] Shklar *Ordinary Vices,* 22.

[144] See Hall, "Contingency, Confidence, and Liberalism in the Political Thought of Bernard Williams." for this observation.

practices in a way that makes social and psychological sense. But we may not be able to find anything that will meet a demand for justification made by something outside those practices."[145]

We can feasibly imagine convincing someone to incorporate moral considerations into her deliberations, to weigh certain moral considerations more heavily, or to represent some as having a "virtually absolute priority."[146] It would be misleading, however, to accuse those who do not accept the system, or who do not accept it absolutely, of moral wrongness. We can only conclude they have not (yet) been sufficiently persuaded. In most cases, the best we can hope for is that people will adhere to degrees of law or morality. I believe this is the general idea behind the political realism of Shklar and Williams, who recognize both the importance of and the limits to these systems. In *Ordinary Vices*, for example, Shklar claims that law acts an alternative to violence, especially the institutional kind with civil rights being the paradigm.[147] In *Ethics and the Limits of Philosophy*, Williams highlights the importance of obligation, since it ensures reliability, and offers a deliberative, practical conclusion.[148] The danger only arises when we characterize moral or legal considerations as the only considerations that matter or that should matter in political deliberation. My point, then, is that modeling politics on the grounds of moralism or legalism seems too much of a departure from our world, and so, cannot really make sense to us. When philosophy misunderstands justice and obligation as the entire world of politics rather than a feature of that world, we misuse those concepts. We may then grow willfully ignorant of the kinds of legal and moral cruelties that stem from such uncompromising social ideals.

For many, accepting politics as a process of persuasion exposes us to too many risks and too much uncertainty to provide any normative grip. What about the demagogue who preys on our fears? What about our unchecked biases that make us predisposed to be more persuaded by certain arguments and not by others? The realist does not pretend there are simple answers to these questions but instead affirms that ideal theory has failed to deliver the certainty it promises in part because it has refused to admit that politics has always been a matter of persuasion. The

[145] Williams, *Ethics and the Limits of Philosophy*, 114; 185.

[146] Williams, *Ethics and the Limits of Philosophy*, 185.

[147] Shklar, *Ordinary Vices,* 214; 237.

[148] Williams, *Ethics and the Limits of Philosophy*, 174-196.

point here is that we deceive ourselves by thinking politics can be anything more. Once again, admitting this does not amount to losing normativity or repudiating our convictions without serious thought. For Williams, the aim of persuasion "is to encourage some forms of it rather than others. This is not a technical task, like clearing a radio of static. It is a practical and ethical task, like deciding who can speak, how and when."[149] The aim of political realism, then, is to clarify the practical limits of ideal theory and to suggest additional ideals for politics, such as preventing the worst forms of harm that individuals can inflict on one another. Which system of rules and principles we agree to, how we agree to them, and how we go on to justify them to others is always a political practice. Abstracting normative principles from the practices which bring them about leads to unhelpful and sometimes cruel conclusions. The point of political philosophy then, is to avoid those cruel conclusions.

6. Critical Realism

In Part I, I showed that Rawls's method does not simply rest on a set of evaluative principles suitable for a theory of justice but puts forward a model for justifying those principles. This model, I observed, is realistically utopian insofar as it sets out a process of justification with human deficiencies in view and refuses to appeal to the self-evidence of general concepts or convictions.[150] This is what Rawls means when he asserts principles must adhere to a political and not a metaphysical justification. I then argued that, in his attempt to control for the contingencies of practical identity, Rawls provides a moral answer to the question of legitimate power and so fails to abide by his own standards. Rawls's model does this by first controlling for all the contingencies that bias our moral identity. He then justifies the two principles according to a shared moral identity and finally applies these principles to politics. Williams calls this method "political moralism," which in light of Shklar's work, we can understand as an example of rule-oriented thinking.[151]

Both Shklar and Williams are primarily critical of how Rawls's political moralism characterizes the conditions for the possibility of normative justification. Within Rawls's model,

[149] Bernard Williams, *Making Sense of Humanity* (Harvard University Press, 1995), 148.

[150] Rawls, *A Theory of Justice*, 19.

[151] Williams, "Realism and Moralism," 2.

legitimate power is not justified against a backdrop of "a mere *modus vivendi*," as Rawls puts it in *Political Liberalism*, but emerges from an objective source of normativity, one that lies deeper than any empirical source.[152] Rawls provides a "moral answer", according to Williams, insofar as he locates this source in moral identity (i.e., the noumenal self). Rawls's model turns to moral identity because it presumes that practical, contingent identities are only capable of justifying so much. For political moralists, there must be some solid ground from which contingency springs, some rational principle from which ethical value is realized.[153] The normative force behind our practical identities, therefore, stems from the value we place on the practical necessity of our shared moral identity, without which we lose all sense of justification.

I hope to have demonstrated the methodological problems associated with using moral reasoning to justify political principles, the most pressing of which is that moral identity is characterized as external to political life, which suggests a distance from and devaluation of empirically informed justifications. Though Rawls's model is methodologically impressive, it ultimately provides a moral justification for his normative principles. In other words, Rawls's model does not provide a justification of how moral considerations are brought to bear on political life in a way that makes social and psychological sense.

The cruelty principle is a political rather than a moral or legal principle, and accordingly, our actions do not treat it as a rule-bound procedure. Putting cruelty first may take virtually absolute priority, but this does not imply that our actions are never without exception or doubt. When we put cruelty first, we express some normative standard or course of action for political life, one that remains vigilant against the state's coercive power. Historically and presently, we have seen the abuse of authority and the sheltering behind rules from official agents at every level of government and we are well aware that such acts are felt most by the poor and the powerless.[154] In light of these "undeniable actualities," Shklar calls for the institutionalization of suspicion, which simultaneously looks to law as the apparatus for reducing cruelty and reminds us that "the

[152] Rawls, *Political Liberalism*; quoted in Williams, "Realism and Moralism," 2.

[153] For this reading of practical identifies, see Christine Korsgaard, *The Sources of Normativity* (Cambridge University Press, 1996).

[154] Shklar, "The Liberalism of Fear," 28.

power to govern is the power to inflict fear and cruelty and that no amount of benevolence can ever suffice to protect an unarmed population against them."[155]

Consider here two final objections to my argument that a realistic approach to political principles ought to orient us towards the cruelty principle.[156] The first objection has to do with the concern that realism leaves us with little hope and wonders whether realist principles offer any valuable insight for a progressive political project, such as what can be done in a constitutionally stable democracy.[157] I respond by noting that a *summum malum* orients towards certain normative ideals, it need not stand as the primary aim of liberalism once and for all. Still, even if we claim that the liberalism of fear is not the totality of politics or political aspirations there remain, in Williams's view, "questions of what should be attended to second, when what must be attended to has been, at least to a decent extent locally attended to."[158] Neither Shklar nor Williams have much to say about how we might progress from the liberalism of fear to something more aspirational, perhaps because they believe it would simply distract us from what we need here and now. Shklar asserts that the liberalism of fear is a first principle, not the only principle. Realism is tasked with constructing more ambitious political theories, such as egalitarianism and dignitarianism, while remaining firmly in the world.

This leads me to the second related objection, which asks why we cannot simultaneously pursue a *summum malum* and a *summum bonum*? When the choice is presented as a dichotomy, the decision is quite clear. For example, if compelled to choose between a policy aimed at improving prison conditions by reducing overcrowding and ensuring access to adequate healthcare and nutrition or providing comprehensive prison college education programs, the cruelty principle would lead us towards the first option. One justification for this choice might highlight the difficulty one would have in pursuing an education at all if one's basic needs are not first secured. And yet, if we are realistic, we know that policy is seldom presented as a

[155] Shklar*, Ordinary Vices*, 238.

[156] See Andrea Sangiovanni, "Normative Political Theory: A Flight from Reality?" in *Political Thought and International Relations*, ed. Duncan Bell (Oxford University Press, 2009), 219-239.

[157] Sangiovanni, "Normative Political Theory: A Flight from Reality?"; see also Edward Hall and Dimitrios Tsarapatsanis, "Human Rights, Legitimacy, Political Judgement," *Res Publica* 27 (2020), 184.

[158] Williams, "The Liberalism of Fear," 60.

dichotomy, and we can reasonably envision a scenario in which both projects are pursued.[159] If this is possible, then why abandon the higher reaching political values?

In general, I agree with the sentiment underlying these concerns and with the desire to think in aspirational terms. I also follow Williams in recognizing that what is most basic in human terms may differ from what we need or desire in political or social terms.[160] Here, I wish to affirm the importance of such ideals, particularly if they help secure the foundational conditions for freedom and equality. Nonetheless, I argue that any higher reaching principle must be constructed in light of what has been locally secured and must be articulated and justified from within political life.

What underlies the concern regarding realism's ambivalence towards progressivist utopianism is a broader worry that it is insufficiently ambitious and so fails to offer enough hope for a normative political project. Although it is true that realists lower the theoretical standards for political ideals, they never abandon practical commitment. If we were to truthfully express and embody the cruelty principle, a great deal of hope remains. Still, I think it is important to recognize the distinct challenge realism faces, which is best articulated by Andrea Sangiovanni:

> [H]ow can a conception of a political value be sufficiently *critical*, given the more historical, context-bound interpretation role that it is intended to play? The challenge here is to explain how practice-dependent views can avoid arbitrarily favouring the status quo. Should not the articulation of a value like justice, after all, help us to 'get outside' our current institutions and practices, and to evaluate them from a point of view free from the historical injustices congealed in existing forms?[161]

---

[159] In addition to thinking about rights in terms of what has locally been secured, realists consider what can actually be achieved in that locality. So, although we can conceive of one set of policies that aim to prevent the worst and another aimed to achieve the best, realists remind us about the cost of rights. The fact that rights have cost, often staggeringly so, is easily overlooked. Faced with these kinds of considerations, then, it makes sense to divert more resources towards preventing the worst. Part of the reason why the choice between a *summum bonum* and a *summum malum* is so difficult, both philosophically and politically, is because we must first consider whether that more aspirational ideal is actually reasonably counted as a condition that might lead to cruelty and fear. For example, the right to work may not seem inherently connected to cruelty, but one can easily imagine how a society where only some citizens are able to obtain work, might lead to rather cruel conditions. The denial of work may itself be cruel. See Colin Farrell, "Justice in Ideal Theory: A Refutation," *Political Studies* 55, no. 4 (2007), 844-864.

[160] Williams, "The Liberalism of Fear," 60.

[161] Sangiovanni, "Normative Political Theory: A Flight from Reality?," 235.

I agree with Sangiovanni that realism should not arbitrarily favor the status quo and that part of what it means to become more realistic and more truthful about politics is to become more critical. I disagree, however, that a critical approach is one that is free of history. It seems to me that Sangiovanni seeks an ethical theory that enables us to observe and reflect on our present circumstances from an external standpoint. According to Sangiovanni, reflection from an external standpoint is best positioned to expose our ethical beliefs and practices as prejudiced. This strikes me as an impossible task since reflection through an external standpoint presupposes that we can extract ourselves from the very practices we are trying to examine.[162] To be sure, we can—and ought to—reflect on our ethical beliefs to ensure we are not arbitrarily favoring the status quo. Yet, such reflection will have more to do with revealing that what we have taken as self-evident or obvious might actually rest on self-deception or wishful thinking. Once we recognize the historical contingency of our practices, we might be motivated to change them or improve them. Alternatively, the reflection will remind us that what we have taken for granted has been precariously secured and that we must fight to hold onto the status quo. Either way, this process cannot succeed without a more developed historical sense, which I argue, could help us become more truthful about our beliefs and practices.

[162] See Williams, "Theory and Prejudice," in *Ethics and the Limits of Philosophy*: 93-119.

III.

FEMINIST POLITICAL REALISM

IN ANDERSON AND YOUNG

One of the most significant contributions feminist philosophy has made to meta- and normative ethics involves what is often called "nonideal theory."[163] Neither uniquely feminist nor exhaustive of feminist approaches, nonideal theory has carved out a distinct place within feminist philosophy and normative theorizing more broadly.[164] Despite its influence, however, feminist nonideal theory lacks a clearly articulated methodology. This chapter addresses that gap by identifying the common methodological traits of feminist nonideal theory, organized around three main components: a metaethical framework, a negative aim focused on ameliorating oppression, and a positive ideal of individual and collective empowerment.

By identifying these three common traits, I also redefine the method as feminist political realism, or feminist realism for short. I argue that there are good reasons to prefer this label over feminist nonideal theory. First, the latter suggests a merely oppositional and parasitic relation to Rawlsian ideal theory. As I will argue, feminist realists evince positive metaethical and normative commitments that the nonideal label belies. Second, situating feminist philosophers, and especially feminist egalitarians, within political realism benefits realism itself. We encounter politics in prosaic spaces such as in the home, the workplace, and in our personal relationships. Political realism needs the theoretical and strategic resources we find in feminist philosophy.[165] A feminist-informed perspective can also correct what J. Ann Tickner calls the "masculine bias" in political realism.[166] Finally, describing nonideal feminist egalitarians as political realists helps

[163] See Amia Srinivasan, "Feminism and Metaethics," in *The Routledge Handbook of Metaethics*, ed. by Tristram McPherson and David Plunkett (Routledge, 2017), 595-608.

[164] For a comprehensive overview of nonideal theory, see Lara Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map," *Philosophy Compass* 7 (9), 654-664.

[165] Elizabeth Frazer, "Including the Uncountable: The Realism of Feminist Politics," *Juncture* 21, no.4 (2015), 315-318.

[166] J. Ann Ticker, "Hans Morgenthau's Principles of Political Realism: A Feminist Reformulation 1988," in *A Feminist Voyage Through International Relations* (Oxford University Press, 2014).

develop a more substantive and coherent account of liberal political realism. Disagreement will, of course, persist within this framework. My aim is therefore to articulate an account that remains open to revision in the future.

At the core of a feminist realism method is a concern with how we might ethically challenge our present values and practices that disadvantage some while advantaging others. It rests on two general assumptions. First, individuals do not autonomously shape the values they hold, and even less so the attitudes and practices that render those values intelligible and appropriately expressed. Second, attitudes and practices structure social relations by defining one's relational status and by shaping how individuals are expected to act and be treated in virtue of their group membership. In this way, practices constitute individuals as members of social groups and define forms of differential treatment.

This relationship between values, practices, and social groups gives rise to two problems for feminist realism. The first problem concerns the frequent misalignment between personal and social values, and between one's values and her socially recognized expressive practices. When such misalignment occurs, individuals may lack intelligible normative avenues through which to express their values. In such cases, one's personal values are frustrated or unrealized. The second problem concerns how normative expectations track group membership. It is not inherently problematic that groups are subject to different expectations. Under conditions of structural inequality, however, a difference in status does not simply invite asymmetrical expectations, which afford greater authority to some groups and greater authority of some groups over another.[167]

When a disadvantaged group comes to question an extant value and its expressive practices, the group will demand a justification from those whom it holds accountable. In the absence of a justification that makes normative sense to those subject to the practices, the disadvantaged group will legitimately call for a practice's revision or abandonment. Under realistic conditions, the path to a justification is not always clear. The demands of the disadvantaged are often silenced or rendered unintelligible to the dominant group. One task of a feminist realist method is to contest values and practices so as to deepen their legitimacy. I argue that this positive aim remains undertheorized. When it is theorized, it is overly idealized and so fails to account for

[167] See Margaret Urban Walker, *Moral Repair: Reconstructing Moral Relations after Wrongdoing* (Cambridge University Press, 2006), 101.

how groups can be formally included into a process of justification but nevertheless excluded from meaningful participation. Drawing on Iris Marion Young's writings, I develop an ideal of individual and collective empowerment, which I argue deepens the legitimacy of our practices under realistic conditions.

This chapter proceeds as follows. To begin, I present a metaethical framework suitable for a feminist realist method. I draw from Elizabeth Anderson's pluralist and expressivist theory of value to account for the manifold ways we value things and the social nature of our valuations. I then provide an account of social groups as shaped by institutional and social practices. After establishing a metaethical framework, I address the negative aim of feminist realism, which is to ameliorate oppression. I define oppression here as harms generated by institutional practices that systemically disadvantage one social group while benefitting another. I then turn to feminist realism's positive aim by expanding on Young's notion of internal exclusions, which speaks to the mechanisms that render individuals incapable of making meaningful demands even when formally included in a justificatory process. I identify two categories: *broad internal exclusions*, which threaten legitimacy by upholding restrictive norms of reasonableness and publicity, and *narrow internal exclusions*, which create epistemic and political barriers to justificatory practices. In contrast to Anderson's ideal of democratic equality, which I argue is too idealized to be fully responsive to these exclusions, I turn to Young's model of individual and collective empowerment, which I argue provides a paradigm for a positive ideal for feminist realism.

1. Pluralist Values and Group Relations

As an attribute of her feminist and pragmatist spirit, Anderson seeks to develop a normative framework to explain and vindicate the pluralism of "ordinary evaluative thought."[168] Central to her theory are three concepts: values, valuations, and evaluations. Values are the things, properties, or ideals that matter to individuals as objects of care, such as persons, activities, and self-conceptions. Valuations are the practical attitudes and emotions through which we express that we care, and involves standards of perception, emotion, deliberation, desire, and conduct.[169] For Anderson, valuations are not simply attitudes towards values, but expressive states. When I

[168] Anderson, *Value in Ethics and Economics* (Harvard University Press, 1993), 3.

[169] Anderson, *Value in Ethics and Economics*, 2.

value someone, I hold a positive attitude toward it and establish a social relationship with them through practices that embody and express my attitude of caring.[170] Finally, evaluation is a process of interpreting the expressive states of our valuations in order to determine whether something is appropriately valued. According to Anderson, "when we wonder whether something is appropriately valued, we wonder whether *we* would be making sense in valuing it."[171] Making sense has two meanings here. First, we ask whether we have good reason for adopting a particular attitude towards our object of care, and second, we ask whether our modes of valuation are intelligible to others.

This practical framework also entails second-order reflection on oneself. This is because, as Anderson suggests, "the grounds of a person's reflectively held values (if she has any) lie in her conceptions of what kind of person she ought to be, what kinds of character, attitudes, concerns, and commitments she should have."[172] Ideals enable us to distinguish between what is important in a personal sense, grounded in who I am and what I am personally attached to, and what is valuable from an impersonal point of view independent of standpoint. Ideals likewise structure the obligations we understand ourselves as bound to. They constitute our practical identities, frame our considerations, and help adjudicate competing commitments.[173] An ideal is obligatory. No trivial or unobjectionable change in my circumstances would warrant abandoning it. If I fail to live up to an ideal or if my ideal is frustrated, I am likely to experience that frustration as humiliating, degrading, or distressing.[174]

Anderson's expressivist and pluralist theory of value illuminates the frequency with which our evaluations of values and ideals come apart from the socially prescribed forms of valuation available. Individuals are not self-sufficient in their expressions of caring. They depend on socially embedded practices to make their values and valuations intelligible to others. But what happens if I lack the appropriate normative avenues to express my valuations? If I disregard the extant practices, my valuation will be unintelligible to others, and I will risk social derision or

[170] Anderson, *Value in Ethics and Economics*, 3; 17.

[171] Anderson, *Value in Ethics and Economics*, 3.

[172] Anderson, *Value in Ethics and Economics*, 6.

[173] Anderson, *Value in Ethics and Economics*, 7.

[174] Anderson, *Value in Ethics and Economics*, 7.

punishment. Nevertheless, for Anderson, the fact that our valuations are socially mediated does not entail acceptance of extant normative practices. What is required, rather, is a procedure for interpreting the social practices that express particular valuations and subjects those practices to critical scrutiny. I will return to this process of justification. However, before arriving there, it is first important to understand how the valuations and practices we find ourselves in shape us into particular people and generate further normative expectations in virtue of our identities.

Differences in normative expectations that generate harms that systematically disadvantage some groups while benefitting others constitute the bedrock of oppression. To make sense of this, feminist realists begin with the normatively relevant fact of social groups. They then assess illegitimate practices by first considering individuals within a social context and second, by examining the social practices that inhere individuals into oppressive, hierarchical social relations in virtue of their group membership.[175] I call this a *politically realist group relations methodology*.[176] There are inherent risks in defining a social group, and yet, one would be hard pressed to discuss an injustice without referring to collections of persons, such as women, African Americans, Native Americans, the poor, the working class, or the disabled, so as to provide description of social groups that offer greater precision to a diagnostic program.

Though a politically realist group relations methodology does not preclude normative individualism, it does challenge political philosophy's tendency to value a group method only insofar as it promotes the interests of individuals.[177] This has the tendency to assimilate or instrumentalize social groups by representing them as either a model of aggregates or a model of associations.[178] The former model specifies a group in respect to some attribute an individual possesses, such as eye color, skin color, or age. Meanwhile, the latter model, represents a collection of individuals within a formally organized institution, such as a university, a political

[175] See Elizabeth Anderson, "Toward a Non-Ideal Relational Methodology for Political Philosophy: Comments on Schwartzman's *Challenging Liberalism*," *Hypatia* 2, no.4 (2009), 130, who refers to this simply as a *group relations methodology*.

[176] Anderson uses this term in reference to Schwartzman's work, though I think it can apply to feminist realism more broadly.

[177] Anderson, "Toward a Non-Ideal Relational Methodology," 132.

[178] Iris Marion Young, *Justice and the Politics of Difference* (Princeton University Press, 1990).

party, or a union.[179] It is clear how associations and aggregations can profoundly shape the trajectory of a person's life, though most philosophers affirm that these groups are ultimately arbitrary from a moral point of view and should therefore offer no significance in themselves nor should they be granted the entitlement to make claims on institutions.[180] Within an individualist methodology, a group is only ever useful if it can achieve ends for the individual.

Feminist realists tend to disagree with models of aggregates and associations because these models ignore the descriptive claim that social groups do exist. Normatively, we may affirm the moral arbitrariness of race and gender though we must admit that descriptively, injustices are overwhelmingly experienced by individuals in virtue of their group membership. Locating group-based injustices does not, in my view, amount to an essentialist, determinist, or reductive conception of identity, nor does it preclude an individualist methodology.[181] Instead, it affirms that the contingencies which lead someone to associate or be associated with a particular group(s) significantly impacts the course of one's life.

Those who deploy a politically realist group method explain the formation and maintenance of social groups along a spectrum that ranges from the intentional to the structural. Young's method asserts the reality of social groups and so adopts a structural account to explain their emergence. According to Young, groups are characterized by their "thrownness: one *finds oneself* as a member of a group, which one experiences as always having been."[182] Our identities are therefore non-voluntary in the sense that we have little control over how others identify us or how we might be associated with "specific attributes, stereotypes, and norms."[183] Though Young acknowledges that one can form an identity in one's own likeness and style, there are particular group ontologies that are near-impossible to escape or join.

Although I largely agree with Young's structuralist description of social groups, I disagree with its stronger ontological commitments. Instead, I follow Ann Cudd, who models groups as simultaneously formed and maintained through intentional conventions, norms, and practices,

---

[179] Young, *Justice and the Politics of Difference*, 44.

[180] Anderson, "Toward a Non-Ideal Relational Methodology," 132.

[181] See Anderson, "Toward a Non-Ideal Relational Methodology," 132-134.

[182] Young, *Justice and the Politics of Difference*, 46.

[183] Young, *Justice and the Politics of Difference*, 46.

but who situates those intentions within a set of historical and social processes that precede any one individual.[184] Cudd writes, "While all action is intentionally guided, many of the constraints within which we act are socially determined beyond the control of the currently acting individual; to put a slogan on it, intentions dynamically interact within social structures."[185] Depending on the social group, the dynamics may favor an intentional explanation over a structural one, or vice-versa. For example, there are near fully-intentional groups, such as the conversational pair, and near fully-structural groups, such as gender. The dynamics between the intentional and structural nevertheless persist. Two educated interlocutors may find themselves unintentionally conversing only with those of a similar educational status; similarly, a person may choose to leave an assigned group and join another, such as in the case of gender identity. Depending on the aim or scope of one's explanation, one may place greater emphasis on either an intentionalist or a structuralist theory of social groups. Since my aim is to explain the basic dynamics of political oppression and since we seldom consider an individual as intentionally choosing their oppression *qua* oppression, my approach to social groups opts for a more structuralist explanation, albeit one that saves room for intentions.

What I find particularly important about the structuralist approach is that it explains how group identity forms absent of internal or intentional states of particular individuals and instead explains how it forms through external factors of the world, including how individuals and institutions treat that person in virtue of an assumed social group identity.[186] While group intentionalism is useful for explaining how a public willingly expresses and embodies voluntary group membership, it faces an explanatory deficit for nonvoluntary social groups where conditions of membership do not rest on intentional states or reflection on those states. For nonvoluntary groups, membership is conferred by individuals sharing in socially significant constraints and are formed and maintained through social treatment. Constraints here are minimally normative insofar as they guide action by creating some possibilities while limiting others. More generally then, a social constraint is best understood as an orientation towards some practices and norms over others due to a rule-bound system of rewards and punishments one can

[184] Ann Cudd, *Analyzing Oppression* (Oxford University Press, 2006).

[185] Cudd, *Analyzing Oppression*, 36.

[186] Cudd, *Analyzing Oppression,* 36,

reasonably expect from one's actions. In addition, rewards and punishments are meted out by institutions and expressed by individuals within those institutions.

Implied in the name, a nonvoluntary group speaks to the limitations any one individual can have over the norms expected of her group identity and reveals the difficulty of addressing deeply structural social constraints. Part of this may be a result of an individual not realizing her social constraints *as constraints* or not reflecting on how certain practices maintain the social group. Some may come to resent their constraints, and few may be bold enough to reject a group constraint entirely. In most cases, however, individuals will come to see their constraints as expressions of preferences, which can in turn pose normative dangers if those preferences become oppressive. Anderson refers to this problem as "the fact of unreasonable pluralism" and holds that feminists may intervene by redirecting activist efforts away from state action and towards other domains, such as civil society, church, and family.[187] Cudd, however, argues that oppressive preferences are often perfectly reasonable given the choices one is presented with. Rational choice contributes to the maintenance of oppression since one's complicitous preference is simply an adaptive preference, formed non-voluntarily and without one's awareness.[188] With enough treatment as a member of a non-voluntary group, an individual rationally adopts the associated constraints of the group and sees these constraints as her own individual preference wholly independent or unaware of her group.

It is worth acknowledging, however, that there are many nonvoluntary groups comprised of individuals who do affirm their identity, though I would hesitate to call the group strictly voluntary in the absence of a willed unity. Still, awareness of and identification with a nonvoluntary group may provide the hermeneutical resources for understanding one's sense of self, may improve self-worth, or provide a feeling of solidarity with those who share the constraints, such as in the cases of consciousness-raising groups. Once a person is convinced of or becomes aware of her social group, she can engage in collective critique in which the social group might affirm some norms, drop others, and begin to will its own.

I will return to these points later in the chapter. For now, I arrive at a tentative definition: A social group is a collection of persons who identify or are identified with a shared social status,

[187] Anderson, "Toward a Non-Ideal Relational Methodology," 131.

[188] Ann Cudd, "The Paradox of Liberal Feminism: Preference, Rationality, and Oppression," in *Varieties of Feminist Liberalism*, ed. Amy R. Baehr (Rowman & Littlefield, 2006), 9.

who share a common history, and who share set of constraints and enabling conditions on action vis-à-vis institutional treatment that social status produces.[189] This definition captures both nonvoluntary and voluntary groups while also allowing for sufficient internalist group identification within a structural framework. In addition, I believe my description is sufficiently weak to accommodate the changing social demands of a diverse array of groups, and strong enough for a group relations method. My definition also captures the fundamental intra- and inter-relational aspect of social groups in which group membership is conferred through institutionally supported treatment. Having outlined a pluralist and social theory of value and shown how values and their expressive practices shape individuals as members of social groups, I now turn to the ways in which certain practices systematically disadvantage some groups while advantaging others. Understanding the nature of this process, and ameliorating its harms, defines the negative aim of a feminist realism.

2. Oppression

The concept "oppression" has increasingly gained traction in resistance movements since the 1960s, and most feminist egalitarians take the phenomena as a starting point for both its negative and positive political aims. Historically, oppression has been the fear of a tyrannical, despotic ruling group over another. In more recent historical configurations, oppression is largely structural and produced through a diffuse network of everyday practices rather than a dyadic sovereign-subject relation.[190] It exists in prosaic spaces—in schooling, bureaucratic waiting rooms, at one's workplace, or in the confines of one's home—and is carried out by well-meaning individuals who do not think of themselves as oppressors but become so through their unquestioned norms and practices and their adherence to the default assumptions, institutional rules, and collective consequences of perceived misconduct.[191]

Not all individuals will feel oppressed, but throughout the course of one's life, many will feel the acute pains of this harm. Fewer still will have, objectively speaking, actually been oppressed. Likewise, many will come to embrace their oppression, or will unreflectively accept

[189] Young, *Justice and the Politics of Difference*, 44; Cudd, *Analyzing Oppression*, 44.

[190] Young, *Justice and the Politics of Difference*, 41.

[191] Young, *Justice and the Politics of Difference*, 42.

their status as simply the way things are. This suggests an asymmetry between the feeling of oppression and oppression as an objective social fact. Most feminists acknowledge the fact of this asymmetry by placing constraints on which claims to oppression count as objective claims.

There are at least three general payoffs to this approach. First, it provides us with a mechanism to refute claims to oppression that do not intuitively or reasonably seem accurate or sincere. One can easily imagine a member of a fascist paramilitary group categorizing himself as oppressed by the tyranny of a democratic government, and feminists would like to say, with some confidence, that this claim to oppression does not amount to an objective social fact. Second, the constraints allow us convince individuals or groups of their oppression and to empower them to resist. Third, establishing objectivity may guarantee a procedure for groups aiming to justify their claims, for example, in a court of law where outcomes are binding.

Though there is general consensus on the need for objective criteria for oppression, feminists argue over how comprehensive and rigid these criteria should be. There are those like Young who are skeptical towards definitions and who worry that a rigid conceptual constraint will *de novo* dismiss outlier claims or lead to what Elizabeth Martínez calls the "oppression Olympics" where groups dispute whose oppression is more fundamental.[192] In addition, Young worries that strong constraints produce reductive analyses, such as the Marxist theories that collapse all struggles of racism and sexism into class domination.[193]Any definition of oppression must therefore be general and amenable enough to not exclude forms of oppression from the start or alternatively to not assign "causal or moral primacy" to others.[194]

Young's definition is sufficiently flexible though substantive enough to capture most of the dynamics of oppression. For Young, oppressed individuals only share in the common condition of suffering "some inhibition of their ability to develop and exercise their capacities and express their needs, thoughts, and feelings."[195] This alone would generate a rather thin account, however,

---

[192] Angela Davis and Elizabeth Martinínez, "Coalition Building Among People of Color," in *Inscriptions* 7 (1994).

[193] Young, *Justice and the Politics of Difference* 42; Kimberlé Crenshaw intersectional method focuses on the intricacies of intragroup differences and the dynamic production of identities along lines of race, gender, and class. As with Young, Crenshaw argues that this methodological failure to recognize the oppression of some identities (both politically and legally) is primarily structural. See Kimberlé Crenshaw, "Mapping the Margins: Intersectionality, Identity, Politics, and Violence Against Women of Color," *Standford Law Review* 43, no.6 (1991), 1241-1299.

[194] Young, *Justice and the Politics of Difference*, 42.

[195] Young, *Justice and the Politics of Difference*. 40.

her "five faces of oppression"—exploitation, marginalization, powerlessness, cultural imperialism, and violence—delimit the possible spheres from which oppression emerges.[196] Presumably, a claim must refer to at least one condition though this does not imply that the categories are mutually exclusive nor strictly additive.

To make Young's account more substantive, I add an additional constraint that an inhibition only becomes oppressive if it reflects an inhibition relative to another group. To articulate this constraint, I turn to Cudd's definition, which names four conditions that satisfy a claim to oppression:

I. *Harm Condition:* There is a harm that comes out of an institutional practice.
II. *Social Group Condition:* The harm is perpetrated through a social institution or practice on a social group whose identity exists apart from the oppressive harm in (I).
III. *Privileged Group Condition:* There is another social group that benefits from the institutional practice in (I)
IV. *Coercion Condition:* There is an unjustified coercion or force that brings about the harm.[197]

Cudd's definition is more precise because it specifies an irreducible form of oppression. The payoff of greater precision, Cudd argues, is that it lends itself to stronger coalition building and provides greater philosophical rigor to a historically diffuse concept. Cudd's coercion condition is particularly important since it articulates the difference between a group that faces ordinary limitations of capacity and one whose limitations arise from social disadvantage. One difficulty, however, with formulating this condition in terms of coercion is that coercion is typically understood as intentional at some level. For example, the state coerces through its laws, and this coercion ultimately reflects the will of the legal framework, which is itself is subject to popular sovereignty. On my view, the distinctive harms at issue need not be intentionally imposed. They instead emerge through a diffuse network of social practices. Accordingly, I revise the coercion condition into a more general power condition, of which coercion is one subset.

According to Marilyn Frye, institutional constraints come from "an enclosing structure of forces and barriers which tends to the immobilization and reduction of a group or category."[198] An oppressed group thus does not require a correlate oppressive group but only a privileged

[196] Young, *Justice and the Politics of Difference*, 48-63.

[197] Cudd, *Analyzing Oppression*, 25.

[198] Marilyn Frye, "Oppression," *Gender Basics: Feminist Perspectives on Women and Men*, ed. Anne Minas (Wadsworth Publishing Company, 1993), 10-16.

group.[199] At most, the maintenance of oppression requires the privileged group to be indifferent to and tolerant of oppressive practices. This is not to say that there are individual persons who intentionally carry out oppressive practices nor does this mean that certain groups have an interest in tolerating oppression. Mills's term "white ignorance", for instance, speaks to forms of productive ignorance that actively refuse to acknowledge race relations in the United States and elsewhere.[200] For Mills, white ignorance stems from impersonal structures that keep groups ignorant of their privilege. Individuals within these groups go about their lives as unknowing agents of oppression and do not understand themselves as oppressors.[201] To come to terms with this fact would require a fundamental reorientation of identity, a reorientation most are often unprepared to undergo.

3. Inclusive Political Justification

Oppressive values and practices frustrate egalitarian ideals. Oppressive values are also the product of illegitimate power arrangements insofar as they secure compliance through mechanisms of force alone rather than through reasons all would accept. Oppressive practices are illegitimate because unjustified.

When we contest values and their expressive practices, we do so through a process of justification by which we make claims on those whom we hold accountable. For Anderson, practices of justification have a second-person or interpersonal form, which is the perspective we take up when we make and acknowledge claims on one another.[202] In a second-person justification, Anderson writes, "a claim of justice is essentially expressible as a *demand* that a person makes *on* an agent whom the speaker holds accountable. Justification is a matter of vindicating claims on another's conduct. Vindication involves demonstrating that the claims are addressed *to* those properly held substantively responsible for the conduct in question, *by* persons

[199] Young, *Justice and the Politics of Difference*; Cudd, *Analyzing Oppression*.

[200] Charles Mills, "White Ignorance," in *Race and Epistemologies of Ignorance*, ed. Shannon Sullivan and Nancy Tuana (New York: SUNY Press, 2007), 11-38.

[201] Young, *Justice and the Politics of Difference*, 42.

[202] See Elizabeth Anderson, "The Fundamental Disagreement Between Luck Egalitarians and Relational Egalitarians," *Canadian Journal of Philosophy* 36; supplementary volume (2010), 1-23; Anderson, "Toward a Non-Ideal, Relational Methodology for Political Philosophy"; and Stephen Darwall, *The Second-Person Standpoint*, from whom Anderson borrows the idea.

entitled to the moral authority or standing to hold them to account."[203] On this account, a second-person or interpersonal form of justification is well-poised to respond to the problems raised by feminist realism, since claims of injustice contain this interpersonal form, whereby one group makes a demand of equality *to* another group who is held accountable.

In its ideal form, such a justification expresses a moral norm of equal respect for all.[204] Anderson refers to this ideal as *democratic equality*, in which person stand in relations of equality rather than hierarchy and domination.[205] Relational equals, Anderson maintains, are not subject to relationships of domination, exploitation, marginalization, cultural imperialism, or violence.[206] Within this ideal model of justification, deliberative processes in which all are formally included, and in which all participate in relations of equality, are best positioned to undermine illegitimate values and generate more legitimate ones in their place. Along with inclusion and equality, we also expect norms of reasonableness and publicity associated with legitimate justificatory procedures.

The norm of inclusion requires that outcomes are legitimate if all those affected by the decision are included in the process, with "affected" meaning that the decision and policies created by a justification "condition a person's options for action."[207] What inclusion presupposes then, is a norm of political equality, both in the positive sense that all those are not simply included in the justification, but included on equal terms, and also a negative sense that they are free from domination.[208] So, although we can analytically distinguish between inclusion and political equality, the two norms entail one another. As Young notes, "in real political conflict, when political actors and movements protest exclusion and demand greater inclusion, they invariably appeal to ideals of political equality and do not accept token measures of counting people in."[209] When inclusion and political equality obtain, participants feel confident

---

[203] Anderson, "The Fundamental Disagreement Between Luck Egalitarians and Relational Egalitarians," 3.

[204] See Anderson, *Value in Ethics and Economics*, 101.

[205] Elizabeth Anderson, "What Is the Point of Equality?," *Ethics* 109, no. 2 (1999), 313.

[206] Anderson, "What Is the Point of Equality?," 315.

[207] Young, *Inclusion and Democracy* (Oxford University Press, 2000), 23.

[208] Young, *Inclusion and Democracy*, 23.

[209] Young, *Inclusion and Democracy*, 24.

that the decision will be supported by the best reasons rather than "fear or force or false consensus."[210] Confidence can only be maintained, however, when participants have a commitment to reasonableness, which entails dispositions such as the willingness to listen to others, to reach agreement, and to remain open-minded.[211] Through this norm, participants hold one another accountable.[212] Participants make themselves intelligible to others by explaining their demands, interests, and needs in ways that others can understand even when they disagree.

In addition to these norms structuring the deliberation itself, there are also implicit regulative norms about the appropriate modes of expression of inclusion, political equality, reasonableness, and publicity when guiding social co-operation in a way that does not reflect the ordinary vices of political life, such as inconsistency, impartiality, ignorance, double-standards, and pragmatic self-defeat.[213] The model of deliberative democracy assumes certain features of deliberative life, such as participants having ample time and energy for meaningful deliberation and minimal prejudices.[214]

Ideally, when we make claims on one another, we approximate a form of pure procedural justice in which the procedure itself is taken to secure legitimate and objective outcomes. The negative and positive aims of feminist realism therefore dovetail: claims against oppression are, at the same time, claims to moral equality. It conceives of equality as a social relation in which individuals regard one another as equals, appeal to principles acceptable to others, by engage in mutual consultation, reciprocity, and recognition.[215] In this way, democratic equality expresses not merely an ideal of equal distribution, but one of equal recognition. To live in an egalitarian community is "to be free from oppression and to participate in and enjoy the goods of society, and to participate in democratic self-government."[216]

Intuitively, this approach appears unobjectionable, since an outcome reached through

---

[210] Young, *Inclusion and Democracy*, 24.

[211] Young, *Inclusion and Democracy*, 24.

[212] Young, *Inclusion and Democracy*, 25.

[213] Anderson, *Value in Ethics and Economics*, 94.

[214] Young, *Inclusion and Democracy*, 33.

[215] Anderson, "What Is the Point of Equality?," 313.

[216] Anderson, "What Is the Point of Equality?," 313; 315.

rational agreement is preferable to one shrouded in bias and arbitrary decision-making. Under ideal conditions the normative connection between justification and legitimacy is straightforward: a justification that begins from a legitimate starting point yields legitimate results. But actual deliberation too often fails to meet these standards of idealized rationality and of idealized participation. Anderson herself admits this much. To this point, she writes, "it is an open question to what degree human communities are capable of achieving objectivity in judgements or whether an objective agreement reached by a local community can be extended to the entire human community, or to all rational beings."[217] According to Young, the reality of these existing defaults in our present deliberative and justificatory practices produce illegitimate results, in which the structural inequalities reinforce greater inequalities and illegitimate outcomes.

Whether an agreement is legitimate or illegitimate is a matter of degree. What is certain is that our current practices of justification fall woefully short of this ideal, even if they partially embody the norms of inclusion, political equality, reasonableness, and publicity. A present limit of Anderson's account is that it offers little guidance on how these justificatory practices respond to our actual conditions that seek to exclude certain groups from deliberation or that fail to meaningfully listen to their claims. Feminist realism that offers a positive ideal must do so in a way that is responsive to these justificatory shortcomings and in a way that deepens practices of justification in light of them.

4. Broad and Narrow Internal Exclusions

Real justifications often violate norms of inclusion and undermine the legitimacy of their outcomes.[218] The most obvious ways they violate this norm is through what Young calls *external exclusions*, which are practices that intentionally or inadvertently exclude individuals and groups from discussion and decision-making, such as restrictive voting laws. Less noticed, Young argues, are *internal exclusions*, which occur when individuals and groups are formally included in the decision-making process but nevertheless precluded from meaningful participation.[219]

---

[217] Anderson, *Value in Ethics and Economics*, 94.

[218] Young, *Inclusion and Democracy*, 52.

[219] Young, *Inclusion and Democracy*, 55.

When someone is internally excluded, she is not treated with equal merit. Her ideas are dismissed and patronized. Considering that most members in liberal societies are formally included into the decision-making process, internal exclusions present a larger and less recognized challenge to inclusive political justifications. I therefore extend Young's account by identifying two forms of internal exclusions that a feminist realism must address. Although not exhaustive, these forms are particularly harmful to the legitimacy of justificatory outcomes.

In particular, I focus on two problems that presently undermine the legitimacy of political justifications. The first, which I call *broad internal exclusions*, arises when political norms of inclusion are conflated with moral norms. For example, norms of reasonableness are conflated with norms of articulateness and norms of publicity are conflated with agreement or unity. The second problem, which I call *narrow internal exclusions*, concerns the epistemic and political barriers that prevent certain groups from making demands or advancing claims at all. Clarifying these problems, I suggest, enables the development of more realistic and responsive ideals for contesting and deepening the legitimacy of values and practices. Narrow exclusions may be especially pernicious as they prevent social groups from making intelligible demands in the first place, even when formally included in the justificatory process. This, in turn, generates a vicious circle in which structural inequalities inhibit participation and in so doing further diminish the capacity to participate in the future.

Dispositions of Reasonableness, Norms of Articulateness. As I have discussed, justifications take a second-personal form. This requires virtuous dispositions for both parties involved, such as a willingness to listen to others, to remain open-minded, and to reach an agreement. Under realistic conditions, however, norms of reasonableness are conflated with norms of articulateness. To present oneself as reasonable entails dispassionately presenting one's demands. These norms favor highly educated individuals and more importantly dismiss apt expressions of emotion from the start. If a wronged person displays emotions such as anger or frustration, her demeaner will likely come off as unreasonable and her demands for repair will remain unattended to.

It is all too common that victims of oppression face expectations to suppress emotions such as anger for the sake of clear communication. Amia Srinivasan calls these second-order wrongs *affective injustice*, which she defines as "the injustice of having to negotiate between one's apt emotional response to the injustice of one's situation and one's desire to better one's

situation."[220] But even when the aggrieved chooses prudential, political strategy over principled self-defeat, there is no guarantee her demands will be heard. Dispassionate claims of wrongdoing may be equally ineffective at conveying wrongdoing or in generating the empathy required to motivate others towards moral repair.[221] This is because emotions express "reactive attitudes" that can be more effective in conveying one's demands than argumentation alone.[222] Anger and resentment, for example, seek assurance from the offenders that they are responsive to the demands of the aggrieved in part by generating motivational and emotional attitudes such as empathy and shame.[223]

Norms of articulateness also generate the expectation that participants dispositions will lack complex emotional attitudes. Consider partiality, which suggests that moral detachment is the preferred state for model participants.[224] Although impartiality is undoubtedly important, it is equally clear that relations of partiality, such as intimate close relationships teach us about the importance of ethical and moral values and allow us to develop a sense of moral integrity. I draw this insight from Williams, who critiques a conception of moral motivation that invokes a rational application of impartiality and distinguishes it from the non-moral, self-interested point of view.[225] In *Moral Luck*, Williams argues that in a situation in which a man must choose between saving his drowning wife and a stranger, he would be justified in saving his wife in

---

[220] Amia Srinivasan, "The Aptness of Anger," *The Journal of Political Philosophy* 26, no. 2 (2018), 135; See also Myisha Cherry, *The Case of Rage: Why Anger Is Essential to Anti-Racist Struggle* (Oxford University Press, 2021); Marilyn Frye, *The Politics of Reality: Essays in Feminist Theory* (The Crossing Press, 1983); Uma Narayan, *Upheavals of Thought: Intelligence of Emotions* (Cambridge University Press, 2001) for accounts that argue for the aptness, if not the productive force, of anger in politics. See also Martha Nussbaum, *Anger and Forgiveness: Resentment, Generosity, Justice* (Oxford University Press, 2016), who argues anger is counter-productive to the aims of politics.

[221] See Margaret Urban Walker, "What Is Moral Repair?" in *Moral Repair: Reconstructing Moral Relations after Wrongdoing* (Cambridge University Press, 2006), 1-39.

[222] See P.F Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 49 (1963), 187-211.

[223] See Urban Walker, *Moral Repair*, "Resentment and Assurance," in *Moral Repair: Reconstructing Moral Relations after Wrongdoing* (Cambridge University Press, 2006), 110-150.

[224] See also Marilyn Friedman, "The Practice of Partiality," *Ethics* 101, no. 4 (1991), 818-835; and Susan Wolf, *Meaning in Life and Why It Matters* (Princeton University Press, 2012), as well as the care ethics tradition more broadly, for arguments in favor of partiality in ethical and moral life.

[225] See, for example, Bernard Williams and J.J.C. Smart, *Utilitarianism: For and Against* (Cambridge University Press, 1973); "Persons, Character, and Morality," in *Moral Luck* (Cambridge University Press, 1982), and *Ethics and the Limits of Philosophy* (Harvard University Press, 1986).

virtue of that relation alone. In his attempts to remain impartial, say in appealing to some moral principle that would justify his choice, Williams states that the man would have thought "one thought too many."[226] Of course, this does not refute the importance of impartiality in practical reasoning. A sympathetic attachment to partiality does not entail that we are altogether blameless when preferring partial relations over impartial, moral demands. This would be, in Williams's terms, "pathological," since we exist in the world and so face possibilities of conflict.[227] My point is rather that expressing dispositions of partiality does not entail a self-interested point of view nor does it indicate a defect in one's practical reasoning as norms of articulateness would have us believe.

Publicity Through a Common Point of View. Just as participants are expected to put aside their affective responses to wrongdoing for the sake of political communication, they are also expected to put aside their demands for the sake of the common good. For Anderson, the common good is the prior condition for and ultimate aim of a justification. "It is pointless," Anderson claims, "to engage in justification when the two parties have no interest in reaching agreement."[228] What is more, the justification cannot meaningfully begin without the starting point of a common ground. The point of proceeding from a common ground and aiming towards a common point of view is that it moves participants from parochial, self-serving, and subjective perspectives, towards rational, impartial and objective ones. Anderson's reasons are intuitively compelling. We want our ideals to rest not merely on reasons that are appealing to our own particular interests, but on reasons that withstand broader scrutiny and speak to a wider community of concern. A justification will not have legs if the participants are not committed to reaching an agreement or if they seek an agreement that serves only their interests.

In ideal circumstances, it makes good sense to proceed from a common ground. Under realistic conditions with diverse needs and interests, participants seldom enter deliberation with a shared set of mutually accepted reasons. Whose interests are deferred will not be accidental or equally distributed. Those who already possess greater authority within the deliberative process will have more say over which issues are taken off the table for the sake of moving the argument

[226] Williams, *Moral Luck*, 18.

[227] Williams, *Moral Luck*, 17.

[228] Anderson, *Value in Ethics and Economics*, 104.

along. It is often the case that deep conflicts and the refusal to abrogate one's interests and needs stem from a broader context in which groups have been asked to wait their turn for the sake of the common good. With enough time, their conditions may deteriorate or become intolerable. When they refuse to bracket their demands any further, they face admonishment for being difficult or obstinate. The justification will breakdown.

Arguments in favor of the common good ultimately presume that we make ourselves intelligible through commonality, which is itself a controversial claim. Even if mutual identification is feasible, which I doubt it is, there is an additional question of whether we desire it. From another perspective, publicity in politics occurs through the plurality of our identities and demands.[229] Admittedly, plurality makes the task of political communication far more complex and thus more likely to reveal structural conflicts than to produce any lasting common interest. However, by becoming more realistic about the complexity of deliberation and justification, we are better positioned to secure a more legitimate because more inclusive agreement.

Situated Hermeneutical Inequality. Under conditions of structural inequality, there are often situations in which social groups lack the hermeneutical resources they require for making intelligible demands. Miranda Fricker refers to this condition as *situated hermeneutical inequality*, which she defines as a situation in which "the subject is rendered unable to make communicatively intelligible something which it is particularly in his or her interest to be able to render intelligible."[230] Lacking hermeneutical resources is a common enough occurrence for anyone. It rises to an injustice when an individual is systematically affected because she is blocked from access to the interpretive resources—be it epistemic, material, or political—she needs to make intelligible demands on another. In other words, there exists epistemic, material, or political resources that would enable her to articulate the conditions of her disadvantage but is structurally foreclosed from accessing those resources.

It is often the case that individuals lack the shared interpretive resources to make demands

[229] Young, *Inclusion and Democracy*, 111. See also Hannah Arendt's *The Human Condition* (Chicago University Press, 1958), from whom Young borrows the idea of plurality. For Arendt, plurality's character is defined by both equality and distinction. She writes, "If men were not equal, they could neither understand each other and those who came before them…If men were not distinct, each human being distinguished from any other who is, was, or will ever be, they would need neither speech nor action to make themselves understood" (175-76).

[230] Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford University Press, 2007), 161-162.

but are nevertheless aware of some harm occurring. Such is the case with sexual harassment. Prior to the concept's widespread political and legal adoption, certain workplace norms between men and women were structured by forms of power that denigrated women in the workplace. Women often described these norms as inappropriate, unwanted, or wrong, yet lacked a concept that would unify these practices into a systematic and systemic problem of the workplace. When individuals are subjected to social norms to which they have not meaningfully consented to, the consequences extend beyond the immediate harms of personal violation. She might feel unsafe in her workplace and incapable of trusting those around her. The dissonance between her own values and those of the workplace may even come to knock her faith in her own ability to make sense of the world.[231] Once her confidence erodes, her ability to meaningfully participate in practices of justification will erode twice over. Try as she might to grasp for the concept, she cannot find it on her own and, with enough time, she may cease grasping altogether.

Adaptive Preferences and Double Binds. A second narrow internal exclusion comes in the form of adaptive preferences, in which an individual endorses the norms associated with her disadvantaged status as part of her own desires and preferences. She thus becomes a willing agent of her own oppression.[232] As Amartya Sen writes, "there is much evidence in history that acute inequalities often survive precisely by making allies out of the deprived."[233] Adaptive preferences bear resemblance to the notion of "false consciousness" in feminist theory, though I prefer the former term, since false consciousness implies that our desires are products of ideology. Adaptive preferences, by contrast, arise from oppressive conditions that, according to Sandra Bartky, "fastens us to the established order of domination, for the same system which produces false needs also controls the conditions under which such needs can be satisfied."[234] What further distinguishes adaptive preferences from false consciousness, and renders them particularly challenging to a feminist politics, is that an individual can be fully aware of her preference as shaped by oppressive structures and yet still rationally endorse or act on them. For

---

[231] Fricker, *Epistemic Injustice*, 163.

[232] See Serene J. Khader, *Adaptive Preferences and Women's Empowerment*, (Oxford University Press, 2001); Cudd, *Analyzing Oppression*, 179-183.

[233] Amartya Sen, "Gender and Co-operative Conflicts," in *Persistent Inequalities, Women and World Development*, ed. Irene Tinker (Oxford University Press, 1990), 126.

[234] Sandra Bartky, "Narcissism, Femininity, and Alienation" in *Femininity and Domination: Studies in the Phenomenology of Oppression* (Routledge Press, 1990), 42.

example, a woman might decide to remain at home for childcare in circumstances where the family cannot afford daycare and her husband has greater earning power. The agential ambiguity that exists when a person endorses the very conditions that disadvantage her makes for a challenging problem for feminists. When we encounter someone who endorses disadvantaging values, it is unclear whether and to what extent her acceptance is causally or rationally produced.[235] Even then, adopting extant practices is rational under the choice situation she is presented with.

Such is the case in oppressive double binds, which Frye defines as "situations in which options are reduced to a very few and all of them expose one to penalty, censure, or deprivation."[236] Double binds become oppressive when unfairly imposed on individuals in virtue of their group membership, such that a treatment comes to express or embody an inequality of authority, status, or standing. They take the form of an adaptive preference because, as Sukaina Hirji writes, a person's "prudential good is bound up with their ability to resist oppression; double binds are choice situations where no matter what an agent does, they become a mechanism in their own oppression."[237] For example, due to the norms and expectations over "good mothering," women are socialized to care about something or some persons more than their male counterparts. As a result, they are placed in what Susan Moller Okin refers to as "a cycle of socially caused and distinctly asymmetric vulnerability" relative to someone else.[238]

Consider the mother who, owing to her socialized altruistic disposition, comes home from a day of work only to cook, clean, and prepare the children for bed while her husband relaxes on the couch. No one here is forcing her to take on these tasks or at least not in the strong sense in which we think of coercion. Indeed, she adopts these preferences as her own because she desires to be a caring mother, she believes good mothering involves self-sacrifice, or she feels she has no other choice. She may even dissuade her husband's attempts to help because she finds his ability lacking (perhaps she thinks he is a bad cook) or because she would view herself as a bad wife

---

[235] Bernard Williams, "Realism and Moralism in Political Theory," in *In the Beginning Was the Deed*, 6; and Williams, "Truthfulness, Liberalism, Critique," in *Truth and Truthfulness* (Harvard University Press, 2002).

[236] Marily Frye, *The Politics of Reality*, 12.

[237] Sukaina Hirji, "Oppressive Double Binds," *Ethics* 131, no. 4 (2021), 643.

[238] Susan Moller Okin, *Justice, Gender, and the Family* (Oxford University Press, 1989), 138; see also Lavender McKittrick-Sweitzer, "Care Exploitation," *Feminist Philosophy Quarterly* 9, no. 3 (2023), for their analysis of care exploitation.

and mother if she failed to uphold these social norms. The question then is whether the justification these social norms rest on, ones that advantage one group and disadvantage another, is a legitimate one.

Feminists would like to say no; or at the very least, they would desire more legitimate justifications. Their reluctance to accept the justification as it stands rests on the intuition that even though someone endorses disadvantaging norms, the choice itself has been shaped and produced by an illegitimate power. In other words, the disadvantaged accept the supposed legitimation because the techniques of belief formations are themselves mechanisms of unmediated or unendorsed power.[239] In this case, what is producing her preference formation is an institutional structure of rewards and punishments relative to her ability to fulfill or her failure to fulfill expected social norms.

5. Ideals of Individual and Collective Empowerment

What I have shown in the previous section are four prevalent sources of exclusion that undermine the legitimacy of justificatory practices and the values and social norms that result from them. These are not meant to be exhaustive examples of internal exclusions, though they point to just some of the more pervasive forms. All four concern internal exclusions of justification that occur under realistic conditions. The broader forms of internal exclusion arise when the norms meant to secure legitimate justification distort the norms governing justifications. Dominant norms of articulateness replace dispositions of reasonableness and thereby render the demands of those who do not speak in the dominant style less effective. Similarly, under the norms of publicity, one is expected to make oneself intelligible by beginning on common ground and assuming a common point of view. The narrower forms of internal exclusion prevent individuals from making intelligible claims in the first place, even when formally included. Situated hermeneutical inequality prevents individuals from understanding their experiences of wrongdoing as systemic rather than individual by foreclosing the resources that would enable them to do so. Adaptive preferences lead individuals to accept or even prefer disadvantages, even when that acceptance is produced by illegitimate means.

One way we can make broad internal exclusions more inclusive is by expanding the norms that structure the justification itself, for example, by expanding acceptable rhetorical styles, such

[239] Williams, *Truth and Truthfulness*, 230.

as emotionally inflected forms of speech. We might also incorporate narrative devices as supplements to our arguments, to convey values in the absence of shared understanding under conditions of plurality. Young has considered these more inclusive forms of political communication, though has devoted less attention to overcoming problems related to narrow internal exclusions. In my view, generating a positive response to narrow internal exclusions ought to be the aim of feminist realism because they tend to generate the political inertia that sustains illegitimate values. The question, then, is what alternatives exist for bringing about a more legitimate justification in light of these narrow internal exclusions?

The challenge of narrow internal exclusions is that they generate political inertia by making social norms appear inevitable, natural, or necessary. A woman in a workplace may feel uncomfortable by her boss's advances or may know his advances are wrong but nevertheless accept this treatment as a natural feature of gender relations. Such belief is expressed in the common refrain "boys will be boys." In cases of adaptive preferences, individuals not only view these norms as natural or inevitable but will desire them.

One way to overcome this inertia is by creating conditions under which individuals and groups learn to express demands to those who they hold accountable. We can understand this process as a model of individual and collective empowerment. Empowerment conveys many senses in feminism and beyond. My account largely draws from Young, who defines empowerment as "the development of a sense of collective influence over the social conditions of one's life," which includes both personal and collective empowerment.[240] Thus understood, empowerment names the process in which individuals engage in dialogue with one another, make sense of the social sources of their disadvantaged status, and realize the possibility of collective action in changing the social norms that contribute to their deprivation.[241] This process transforms participants on an individual level, but within a context of "reciprocal aiding of others."[242] This is because, according to Margaret Urban Walker, a group provides a "community of moral judgement" as well as "a reference and rallying point for our refusal to accept vicious or

[240] Iris Marion Young, "Punishment, Treatment, Empowerment: Three Approaches to Policy for Pregnant Addicts," *Feminist Studies* 20, no. 1(1994), 48.

[241] Young, "Punishment, Treatment, Empowerment: Three Approaches to Policy for Pregnant Addicts," 50.

[242] Young, "Punishment, Treatment, Empowerment: Three Approaches to Policy for Pregnant Addicts," 50.

demeaning norms and the demeaning treatments they justify."[243]

The paradigm of individual and collective empowerment is that of the consciousness raising group. Through the sharing of personal experiences, individuals within the group critically reflect on "the situated social basis of individual action."[244] What initially presents as an individual misfortune or an inevitable state of affairs is rearticulated as an instance of a larger structural phenomenon. In this way, the group functions as a cite of collective interpretation in which participants develop the conceptual and normative resources required to name and criticize the social norms that structure their experience. Practices that once appeared inevitable are denaturalized, thus opening up the possibly for new practices to come through.

This ideal of empowerment closely parallels Anderson's ideal of democratic equality, particularly with its emphasis on reciprocal justification, self-correction, and the shared authority of participants in normative inquiry. Indeed, we can understand individual and collective empowerment as an intermediary ideal that moves us closer to Anderson's more morally ambitious project of democratic equality. In this way, the mode of empowerment mitigates the central question facing feminist realism: how do we generate normatively compelling and legitimate ideals without abstracting away from existing relations of domination and oppression? Although a fuller account of the relationship between collective empowerment and democratic equality remains necessary, Young's framework provides a strong starting point for opening up a positive orientation for feminist realism. It is an ideal that offers resources for normative critique all while deepening the legitimacy of our values and practices.

6. From Social Critique to Genealogical Critique

In this chapter I clarified common traits in a feminist realist methodology. I outlined a metaethical framework appropriate to both the negative aim of feminism (the amelioration of oppressive conditions) and its positive aim (a model of democratic equality). Under ideal conditions, the negative and positive aims of feminism are one and the same: a claim against oppression is a claim to equal treatment and a claim to equal treatment means that one cannot be oppressed. However, under realistic conditions, we are likely to encounter barriers to appropriate

[243] Margaret Urban Walker, *Moral Repair*, 103.

[244] Young, "Punishment, Treatment, Empowerment: Three Approaches to Policy for Pregnant Addicts," 50.

normative expressions of democratic equality as a result of internal exclusions, both broad and narrow. The reality of internal exclusions, and their ability to produce a rift between conditions of oppression and democratic equality, calls for an intermediary ideal of individual and collective empowerment.

The paradigm of individual and collective empowerment is the consciousness-raising group insofar as it provides a clear and practical way of overcoming internal exclusions and the rift they create between the negative and positive aim of a feminist realist project. Yet it remains to be clearly stated what role philosophy can play in supplementing praxis with intellectual rigor. What, in other words, can philosophy do in aiding in the process of collective criticism by clarifying the authoritative and structural sources of our practices and by helping us determine whether there are good reasons to affirm or reject the values and practices that govern our shared life?

In Chapters Four and Five, I argue that philosophy, and more specifically, genealogy, can aid in a process of collective criticism by clarifying the authoritative sources of ethical practices. Whether and to what extent genealogy is itself reason-giving enough to vindicate or subvert those practices is a second question I take up. My primary aim, however, is to demonstrate that political realism requires genealogy, and a broader historical sensibility in order to make sense of the normative grip that practices have over us. Only then can we subject our practices to critical scrutiny

IV.

ON WHAT EXPLAINS:

STORIES, MODELS, AND PARADIGMS

In the two previous chapters, I discussed the methodological dimensions of idealized and realistic justifications. I now follow a similar route through genealogical explanation. Political realism holds that our normative ideals and procedures should reflect the historical and social complexity of political agents and institutions. Yet practice-dependent ideals raise a concern: ideals too closely tied to existing practices will arbitrarily endorse the status quo. Ideals will reflect mere complacency and self-defeat because they lack the critical distance we require for imagining alternative practices. In response to the skeptic, we might wonder whether the problem lies not in practice-dependent ideals as such, but rather in an uncritical approach to practices. The appropriate response to the skeptic, I argue, is not more idealization, but a more critical and empirically informed account of the practices themselves. Explanatory accounts bring into view the historical emergence of practices that once seemed self-evident by revealing their historical emergence and development. Far from mandating the uncritical acceptance of practices, explanations make for better informed justifications because they enable us to assess which practices merit endorsement, which require revision, and which we should abandon. It is in this sense that genealogy becomes indispensable for political philosophy.

There is no single or definitive genealogical method. Since Friedrich Nietzsche. genealogical inquiry has taken on a variety of forms with Foucault and Williams serving as two influential and divergent exemplars. From their respective approaches have emerged further genealogical methods, each one shaped by its distinct philosophical aims. Considering the varied approaches, we might wonder whether some genealogies offer greater explanatory power over others and, if so, whether there is some way of choosing between competing genealogical methods. This question is especially pressing in light of recent interpretations of Williams's genealogical method, such as ones advanced by Miranda Fricker and Matthieu Queloz. I call their respective methods *idealized functionalist models*. According to Fricker and Queloz, idealized functionalist models rival historical and empirically informed genealogies. I disagree

with this claim and argue that idealized functionalist models plunder historical sense from both Williams's genealogy of truthfulness and genealogy more broadly.

What counts as a good or sufficient explanation is largely dependent on the aim of one's inquiry. Newtonian mechanics offers a good explanation of the laws of motion but cannot make sense of why Margaret told a lie. The aim of my inquiry is to provide a politically realist explanation of the emergence and development of an ethical practice, which demands that my explanation be accurate, which is to say, empirically adequate. I refer to empirical adequacy as a cognitive value, which means it is a value that is conducive to producing some truth about the world. I argue empirical adequacy is satisfied through four criteria: determinacy, impartiality, contingency, and regularity. Taken together, these characteristics establish the standards for a politically realist explanation.

With the aim of my inquiry and scope of my criteria set out, I argue that idealized functionalist models lack the explanatory power insofar as they fail to satisfy these four criteria. By contrast, I argue that Williams succeeds on these fronts. The argument proceeds as follows. I begin with an exposition of Williams's vindicatory genealogy and demonstrate the central role of empirical history within his explanatory framework before describing the common traits of idealized functionalist models. I then argue that politically realist explanations should be empirically adequate and specify four criteria for assessing empirical adequacy. These criteria are not intended to be exhaustive, rather, they are individually significant and collectively necessary for a politically realist inquiry. From here I argue that a vindicatory genealogy provides greater explanatory power than idealized functionalist models. Finally, I introduce two further non-cognitive values, freedom and complexity of interaction, that bear on a politically realist explanation. In light of these values, I argue in the following chapter that Foucault's genealogical method offers a more suitable explanatory framework when compared against Williams.

1. Genealogy and the State of Nature

Williams defines genealogy as a "narrative that tries to explain a cultural phenomenon by describing a way in which it came about, or could have come about, or might be imagined to have come about."[245] The particular cultural phenomenon Williams's own genealogy explains is

[245] Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton University Press, 2002), 20.

the value of truthfulness; or, more precisely, it explains how we came to intrinsically value the virtues of truth, comprised of accuracy and sincerity. Williams here does not intend the idea of intrinsic value in the robustly metaphysical way the term frequently invites. Rather, it is related to the idea of importance, an idea he briefly introduces in his much earlier *Ethics and the Limits of Philosophy*. Much like intrinsic value, Williams claims "It is not at all clear what it is for something to be, simply, important. It does not mean that it is important to the universe: in that sense, nothing is important. Nor does it mean that it is as a matter of fact something that most human beings find important; nor that it is something people ought to find important."[246] Rather, when we intrinsically value something, we rank it highly among our considerations and grant it deliberative priority, often "virtually absolute priority."[247] When I intrinsically value truthfulness, for example, I view it as simply, important. In most deliberations, I do not even think to lie or deceive.[248] For Williams, importance and deliberative priority are connected through the idea of obligation, an idea "grounded in the basic issue of what people should be able to rely on."[249] With regard to truthfulness, people not only require information, but need to trust that others will not lie. Ethical life can ensure this need for reliability "by encouraging certain motivations."[250] The importance of truthfulness, its intrinsic value, most accurately reflects the shape of our motivations rather than the property of truth.

Williams's methodological approach blends elements of Hume and Nietzsche's respective genealogies.[251] Hume makes psychological and social sense of our collective reasons for action by reference to motivations we already possess. Nietzsche provides a functional account of collective motivations that resist being understood as such.[252] Like Hume, Williams grounds his story in psychological motivations most agree we already possess, and like Nietzsche, he offers a

---

[246] Bernard Williams, *Ethics and the Limits of Philosophy* (Harvard University Press, 1985), 182.

[247] Williams, *Ethics and the Limits of Philosophy*, 185.

[248] For a similar interpretation of Williams's use of intrinsic value, see Barry Allen, "Another New Nietzsche," *History and Theory* 42, no. 3 (2003), 363-377.

[249] Williams, *Ethics and the Limits of Philosophy*, 185.

[250] Williams, *Ethics and the Limits of Philosophy*, 185.

[251] See Miranda Fricker, "Williams' 'Philosophical Anthropology': A Humean and Nietzschean Synergy," *European Journal of Philosophy* 24, no. 1 (2025), 229-247.

[252] Williams, *Truth and Truthfulness*, 38.

functional account of motivations that have largely been "forgotten." Williams explains motivations to truthfulness in terms of a naturalistic explanation by situating ethical motivation within a network of other psychological dispositions. Naturalistic explanations are not, for Williams, crudely reductive. They do not, for example, explain our motivations in terms of biological evolution.[253]

When I explain why Margaret chose to tell the truth, I could do so by giving an account of her moral obligation to tell the truth. Telling the truth in this context is required as a way of respecting fellow rational human beings, of treating them as an end-in-themselves. Margaret may certainly feel motivated to tell the truth for these reasons, but there will likely be many other reasons. Perhaps she wants to be perceived by others as an honest person or wishes to avoid the shameful reactive attitudes at having her lies uncovered. Telling the truth can also be self-serving. In Jean-Luc Godard's *Breathless*, the protagonist, Michel Poiccard, insists that "Lying is stupid. It's like with poker. It is better to tell the truth. The others think you're bluffing, and that's how you win."[254] There are many reasons to tell the truth. Some stem from ethical motivation and others from different motivations. An accurate explanation will make sense of all, or at least most, of those reasons.

Williams begins his genealogy with what he calls a State of Nature story, rather than with a model.[255] Models are deliberate idealizations of reality for the purposes of capturing its most basic structure. A story, on the other hand, is a narrative about some fictional occurrence for the purposes of illuminating something about our experience. Williams's state of nature story intends to make sense of the phenomenon of truth and truth-telling. Individuals within Williams's state of nature share a common language but lack elaborate technology or writing.[256] They exist in space and time though their conceptions of just what that means is rather vague. They are also generally reflective creatures and are capable of concealing information and deceiving others. Most importantly, these creatures are fictional. The State of Nature is not, as

---

[253] Williams, *Truth and Truthfulness*, 24.

[254] *Breathless (À bout de souffle*), directed by Jean-Luc Goddard (1960, Criterion Collection, 2007), DVD.

[255] Williams draws inspiration for his State of Nature story from Edward Craig's account of the concept of knowledge; see Edward Craig, *Knowledge and the State of Nature: An Essay in Conceptual Synthesis* (Oxford University Press, 1990).

[256] Williams, *Truth and Truthfulness*, 41-57.

Williams puts it, "the Pleistocene," which is to say that it does not tell us something about our hominid ancestors.[257] It is a fiction insofar as it is factually defective in tracing the development of motivations to truthfulness, but not law-defective, which enables Williams to illuminate structural features of our motivations that we might share with these fictional creatures.[258]

The point of Williams's imaginary genealogy is to make sense of human motivations and practices in terms of their functional response to basic needs.[259] Williams writes, "the explanation of the function is unmysterious, because in particular it does not appeal to intentions or deliberations or (in this respect) already purposive thought; and the motivations that are invoked in the explanation are ones that are agreed to exist anyways."[260] Philosophy typically represents reasons for action as purposive and deliberative. But what if our dispositions to the virtues of truth were a result not of ethical reasons but of functional causes? Williams's imaginary genealogy is an attempt to answer this question, without embellishment. The benefit of such a story is that it does not have to appeal much (if at all) to purposive thought since the explanation only calls for an account of our actions as a response to very basic needs. In this sense, Williams's account avoids the problem of idealization raised by Onora O'Neil, who argues that idealized models both omit too much of human psychology that is true and include too much that is false.[261]

Williams constructs his imaginary genealogy from the uncontroversial idea of a *purely positional advantage*, which simply means that humans in the story are not omniscient. A speaker can convey information about a situation because he witnessed it, while the hearer herself did not. As a result of this purely positional advantage, each agent depends on the collective group to acquire information. "They are all," Williams writes, "at an advantage or disadvantage (purely positional, or otherwise) in relation to each other."[262] With this, we have

[257] Williams, *Truth and Truthfulness*, 27.

[258] Williams, *Truth and Truthfulness*, 31.

[259] Williams, *Truth and Truthfulness*, 34.

[260] Bernard Williams, *Truth and Truthfulness; An Essay in Genealogy* (Princeton University Press, 2002), 34.

[261] Onora O'Neill "Abstraction, Idealization, and Ideology in Ethics," in *Moral Philosophy and Contemporary Problems*, ed. J.D.G. Evans (Cambridge University Press: 1987), 56.

[262] Williams, *Truth and Truthfulness*, 43.

the problem of a purely positional advantage and the need for information. In response, the group begins pooling information and so establishes a simplified division of labor. Over time, each agent becomes an observer in various situations and contributes her bits of knowledge—her beliefs—back to the collective pool of information.

One way the group ensures each agent performs the task of acquiring information is by encouraging or rewarding the dispositions that acquire correct beliefs, while discouraging, sanctioning, or shaming those who perform poorly.[263] Truth-acquiring dispositions fall into two kinds, and when combined, they form the virtues of truth. The first, accuracy, is the disposition to seek correct beliefs. The second, sincerity, is the disposition to transport those beliefs back into the pool of information.[264] Accuracy is thus a form of belief and sincerity a form of assertion. Accuracy and sincerity have a value because they are, Williams claims, "useful, indeed essential, to such objectives as the pooling of information, and those objectives are important to almost every human purpose."[265] In other words, in the State of Nature story, individuals are encouraged to cultivate the dispositions of accuracy and sincerity because they are functionally valuable to fulfilling basic needs and purposes we all have.

It is worth noting that in the State of Nature, we are focused on the group's actions, and on dispositions that would benefit the group's needs, rather than on the actions and dispositions of the individuals within the group. From a collective point of view, it makes sense to encourage agents to acquire and express correct beliefs since it benefits everyone to do so. However, at this point in the story, Williams notes a "lack of fit" between the motivation to adopt the virtues of truth for the group, and the motivations for the individuals, who are ultimately the ones expressing them.[266] More specifically, although there are both individual and group reasons to adopt the disposition of accuracy this does not extend to sincerity, since according to Williams, sincerity is a value for other people.[267]

---

[263] Williams, *Truth and Truthfulness*, 44.

[264] Williams, *Truth and Truthfulness*, 44.

[265] Williams, *Truth and Truthfulness*, 57.

[266] Williams, *Truth and Truthfulness*, 58.

[267] Williams, *Truth and Truthfulness*, 59.

Once an agent reflects on the information she has acquired and asks herself whether she shall tell the truth, she may decide to conceal the information or to shape it in a way that deceives others to benefit her own interests. This is because she, like the others in the group, only views sincerity instrumentally, namely as a way of overcoming the purely positional advantage. "Without some help," Williams argues, "the collective value of Sincerity (and in *that* sense, its value to each person) does not translate itself into a reason that each person has for possessing that quality himself."[268] When someone fails to value sincerity for herself, she will not reliably adopt the disposition. If she shares correct beliefs only when doing so aligns with her own interests, then the practice's function is liable to break down. This is not because sincerity only works in conditions of full compliance, but because holding an instrumental value of the disposition generates insecurity among members of the group.[269]

An implicit assumption of Williams's imaginary genealogy, one which I am sympathetic to, is that individuals prefer a situation in which everyone in the group benefits to a situation in which one person wins at the expense of others.[270] If one accepts this premise, it follows that each individual has an interest in seeking assurance that others will express sincerity. For this, Williams turns to a further disposition of trustworthiness to help stabilize the virtues of truth, such that sincerity is understood as trustworthiness in speech.[271] Williams argues a group secures trustworthy behavior by constructing it as an intrinsic value for individuals. So conceived, when an agent asks herself whether she should tell the truth, she views this act as simply, important and ranks it above other considerations in practical deliberation, such as self-interest.

For Williams, one's reasons for regarding certain dispositions as intrinsically valuable lie in how an agent makes sense of their reasons in relation to all their other values. Reasons and motivations conceived from an internal point of view is one of the more disputed elements of Williams's metaethical outlook.[272] Even if the actions of those motivated by intrinsic reasons converge with those motivated by instrumental ones, Williams argues that practices grounded in

[268] Williams, *Truth and Truthfulness*, 58.

[269] Williams, *Truth and Truthfulness*, 89.

[270] Williams, *Truth and Truthfulness*, 89.

[271] Williams, *Truth and Truthfulness*, 94.

[272] See Kate Manne, "Internalism About Reasons: Sad but True?" *Philosophical Studies* 167, no. 1 (2014), 89-117.

merely instrumental reasons, whose force is external, will not reliably endure. This remains the case even if I claim to intrinsically value the practice for instrumental reasons, since those motivations will lack substantive content. If I merely pretend that it is instrumentally useful to treat a practice as intrinsically valuable, my reason for acting will ultimately remain instrumental. The reason may appear external: I may have good reasons to adopt a practice from a collective point of view yet feel no internal motivation to do so. If on the other hand, I regard a practice as intrinsically valuable, my reason is simply that this is the kind of action I do. For Williams, an intrinsic value is strongest when the agent does not even view the value as a value but takes it up and lives it. Perhaps this is why the common phrases "to tell you the truth" or "to be honest" are so jarring to the ear. One should simply tell the truth and not qualify it in an assertion.

Imaginary genealogy demonstrates how the disposition of sincerity benefits the collective group insofar as it aids in co-operation, but it cannot make sense of why a member of that group would feel motivated to adopt that disposition from an internal point of view. To do this, genealogy needs to make sense of truth-acquiring motivations "from the inside."[273] If this were to happen, Williams writes we will have "*constructed* an intrinsic good."[274] In order to do so, the method must make sense of the connection between the collective and individual motivations by moving from the "necessary, structural features" in the imaginary genealogy, and replacing it with its "specific cultural determinations," which we find in history.[275] General reflection demonstrates "that something has to support the disposition of Sincerity, and that an agent should be able to make inner sense of the structure in which Sincerity is embedded."[276] To do so, Williams claims, the abstract account needs "real history," for only this approach will allow us to understand how we encounter the virtues of truth in our own life and among all the other values we hold.[277]

---

[273] Williams, *Truth and Truthfulness*, 91.

[274] Williams, *Truth and Truthfulness*, 92.

[275] Williams, *Truth and Truthfulness*, 93.

[276] Williams, *Truth and Truthfulness*, 93.

[277] Williams, *Truth and Truthfulness*, 93.

2. Three Uses of History in "Truth and Truthfulness"

In *Truth and Truthfulness*, Williams identifies three uses of what he calls "real history."[278]These uses are not optional or auxiliary to his genealogy, but methodologically necessary. Without history, Williams's genealogy cannot make sense of our intrinsic motivations to truthfulness, since it cannot explain how individuals become motivated to adopt sincerity from an internal point of view nor can it justify to individuals why they might adopt the virtues of truth. Although imaginary genealogy demonstrates *that* conditions of trustworthy behavior are required for co-operative communication, it cannot stabilize the virtues of truth, which would involve constructing trustworthiness, understood as sincerity in speech and action, as an intrinsic value. Real history therefore reveals *what* those specific conditions are, *how* they have developed over time, and *where* we can normatively orient an ethical project of truthfulness.

This tripartite distinction—what, how, and where—maps onto the three uses of history. In the first use, history fills in the gaps of imaginary genealogy by explaining the historical values supporting the virtues of truth. It makes accuracy and sincerity thick ethical concepts.[279] In its second use, history extends virtues in the State of Nature story by tracing their development as they go beyond our structural needs. In its third use, history diagnoses by demonstrating how practices have autonomously developed from the State of Nature story such that they no longer serve our structural needs. It therefore illuminates areas where a group might encourage dispositions to accuracy and sincerity.

History Fills In. In the State of Nature story, we only have a sense of the structural requirements for the virtues of truth. These are the conditions that foster trustworthy communication of beliefs and the psychological mechanisms that sustain it. Yet, for us to fully appreciate trustworthiness as an intrinsic value, it is therefore not enough to regard it merely as a basic human need, since a justification on those grounds would only provide an external reason for action. If Margaret fails to see the point in telling the truth, and I tell her that she has good reasons because we, as humans, need it, my justification may not generate sufficient internal normative force to motivate her action. My justification will only generate normative force as an internal reason if the account is integrated with the rest of her motivations, desires, or beliefs.

[278] For a similar interpretation of Williams, see Matthieu Queloz, *The Practical Origins of Ideas* (Oxford University Press, 2021), 172.

[279] For a discussion of thick ethical concepts, see Bernard Williams, *Ethics and the Limits of Philosophy*, 140-148.

What is more, it can make sense of her actions from this internal point of view. It would therefore make little sense to explain Margaret's commitment to truth-telling solely by appealing to basic human needs if such an appeal failed to cohere with the rest of her practical outlook. For this reason, genealogy requires real history to explain why individuals come to adopt the virtues of truth and to make sense of their reasons for doing so. The State of Nature story thus needs history because it needs some descriptive account that would render virtues of truth intelligible as "thick ethical concepts," and situate them alongside other virtues that matter to us.

In Chapter Two, I discussed Williams's idea of thick ethical concepts, which are both descriptive, insofar as they are practice-dependent, and evaluative, insofar as they provide reasons for action. Certain concepts carry evaluative weight. When I say you are truthful, I am not simply stating a fact, like I do when I say you are right-handed. Instead, I am evaluating your actions and character. Most of us wish to be truthful, or at the very least, to be perceived by others in that light, so we do our best to tell the truth. But not just any action will suffice for a thick ethical concept since their descriptive sense entails acting in accordance with a range of practices and corresponding values that correctly map onto a shared understanding of that concept. A concept's intrinsic value, according to Williams, cannot be understood simply in terms of the needs it satisfies, but requires further articulation as a thick ethical concept. It is only by grasping truthfulness in this thicker sense that we come to understand the attitudes, dispositions, and forms of conduct expected of us when we seek to tell the truth. The State of Nature story cannot supply thick ethical concepts, precisely because it does not exist in the world. For this we need real history.

Consider, for example, Williams historical treatment of truthfulness and trust as it relates to sincerity. We typically think of trust as something humans bestow on deserving parties. In the age of Homeric myth, Williams tells us the Greeks established desert through the distinction between friend and enemy. Now, under modernity, we have a more egalitarian conception of desert, one rooted in liberal relations of justice, respect, individual freedom, and the avoidance of manipulation.[280] Everyone deserves the truth, though some deserve more or less of it depending on how they have violated the trust of others. Understood in this way, the reasons that motivate a person to truthfulness is not her need for social cooperation but her respect for another person's freedom, as well as her own. Even in cases where deception seems justified—for instance, lying

[280] Williams, *Truth and Truthfulness*, 122.

to protect someone or to mislead a murderer—we will still experience what Williams calls "ethical regret."[281] Lying injures our integrity and makes us shameful, even when the deceived person does not deserve the truth.[282]

Accuracy also thickens in real history. In the State of Nature story, accuracy simply required that agents acquire true beliefs and presented a simplified account of the division of labor needed for acquiring such knowledge. Under more realistic conditions, agents have greater reflective capacities and knowledge acquisition involves a "conscious choice of policies of investigation."[283] Williams calls this "investigative investment": a price is now set on acquiring information, which includes less tangible dividends, such as the satisfaction of one's curiosity, fame, and respect from peers.[284] Though these motives are self-serving, Williams does not believe they harm truth acquisition provided one earns fame and respect because they have arrived at the truth.[285] In real history, knowledge also comes to express a reflex against self-deception and wishful thinking by adopting the right methods. Some methods are truth-acquiring and others are not. Accuracy now involves having the judgment and courage to know the difference.

History Extends. From our present standpoint, what we demand or expect of truthfulness has extended beyond the structural needs established the State of Nature story.[286] From this realization, one might draw the conclusion that we can release ourselves from the more contingent formulations of a practice and simply adhere to that which all humans structurally, basically need to get on in the world. To dissuade the reader from reaching such a conclusion, Williams's introduces a second use of history, extension, to demonstrate that dropping the contingent manifestations of a value are neither easy nor desirable.

---

[281] Williams, *Truth and Truthfulness*, 115.

[282] Williams, *Truth and Truthfulness*, 120.

[283] Williams, *Truth and Truthfulness*, 123.

[284] Williams, *Truth and Truthfulness*, 124.

[285] Williams, *Truth and Truthfulness*, 142.

[286] Williams's political justificatory story, the Basic Legitimation Demand, draws most heavily on this second use of history. In that story, he first accounts for the structural needs for political legitimacy and then shows how liberalism extends what we find in the Basic Legitimation Demand.

Williams's genealogy deploys history in order to extend the state of nature story beyond our structural needs. Extension demonstrates what we lose were we to abandon our present practices of truthfulness. One extension of accuracy is in the invention of historical time itself; an invention Williams attributes to the Greek historian Thucydides. Prior to Thucydides, truth was not addressed when recounting events in the remote past. When stories reached into the centuries before one's investigation, they rarely referred to a determinate time and place. Williams calls this practice the mode of myth.[287] After Thucydides's invention of historical time, a new way of determinately speaking about the remote past as we do the immediate past was made possible.[288] In recounting Pericles's *Funeral Oration*, Thucydides draws on eyewitness testimony and faithfully represents the events as they were said to have happened. In assessing evidence, he treated the trustworthiness of claims about the remote past as he would claims about yesterday or last week. In so doing, he invented a new way of thinking about our relation to the past, and in particular, what it means to convey historical truth with accuracy and sincerity.

After the invention of historical time, the past is no longer a "a peculiar area in which indeterminate happenings and people could exist."[289] It is real. The claims we make of the past now function as assertions: either there was a time when semi-divine kings walked the earth or there was not. Like in the sciences, historical truth is hard-earned through discovery, research, and luck. Acquiring historical truth demands the right methods, as well as the dispositions that resist vices inimical to truth-acquisition, such as laziness, ambition, and the desire to please one's audience.

Considering the contingent development of historical time and historical truth, does this mean we can, from our present perspective, go without it? Williams argues that we cannot. Whether we adopt the mode of myth or the mode of historical truth when recounting the past is not a matter of taste or preference, but is a demand placed on us by our present expectations of truthfulness. The invention of historical time increased the explanatory power of recounting the remote past, and this is "in terms of *anyone's* conception of explanation."[290] This is not to say

[287] Williams, *Truth and Truthfulness*, 160.

[288] Williams, *Truth and Truthfulness*, 163.

[289] Williams, *Truth and Truthfulness*, 163.

[290] Williams, *Truth and Truthfulness*, 170.

that we could, in some distant future, conceivably go without or lose historical truth from our practices of truthfulness, but only that historical truth is intellectually irreversible: there is no way of expunging historical truth from the historical records. Real history thus exhibits that a practice of truthfulness must now satisfy the explanatory demand of historical truth, even if it exceeds our basic needs.

Just as much as the second use of history demonstrates the importance of practices of truthfulness that extend beyond the State of Nature story, it also constrains what imaginary genealogy can posit as basic human needs. In other words, we can only construct the State of Nature story in light of "what real historical developments there have been."[291] Now that we are so thoroughly embedded in the mode of historical truth, we regard it as self-evident. Williams cautions against viewing our own conception of historical truth as a structural feature of the virtues of truth. As his narrative has shown, truthfulness absent a conception of historical time will not prove "inherently unstable."[292] Though it may be hard for us in the present to make sense of the simplicity with which humans in the State of Nature story report happenings of the remote past, their understanding is not merely an "interesting abstract possibility," but "a matter of historical fact."[293]

History Diagnoses. Not all practices function to fulfill some basic human need, though some bear resemblance to the virtues of truth. If you recall, sincerity is a spontaneous disposition to say what one believes in the State of Nature story. Now in real history, there are forms of spontaneous declarations that are autonomous from sincerity and that serve no clear structural needs. Williams draws from the historical development of authenticity, which like sincerity, is an earnest, spontaneous and uninhibited expression.[294] Unlike sincerity, which is a disposition for others, authenticity is the highest individual ideal of self-understanding and self-constitution. One discloses not the findings of self-discovery, as one does with sincerity, but an uninhibited expression of who one is at a given moment.[295]

---

[291] Williams, *Truth and Truthfulness*, 56.

[292] Williams, *Truth and Truthfulness*, 161.

[293] Williams, *Truth and Truthfulness*, 40.

[294] Williams, *Truth and Truthfulness*, 178; 189.

[295] Williams, *Truth and Truthfulness*, 178; 189.

What can the history of authenticity teach us about the virtues of truth? For one, it can demonstrate what we lose when we have spontaneous disclosure without the virtues of truth; or, conversely, what we gain in remembering accuracy and sincerity's point.[296] In the first use, history "fills in" the gap between the collective need for the virtues of truth and the individual motivation to withhold one's discoveries. In this third use, history reveals a different kind of misalignment: individuals act on their deepest impulses even when doing so comes at the expense of others.[297] In his essay, "Moral Luck," Williams considers Gauguin's choice to abandon his wife and children to pursue an artistic life in Tahiti. Gauguin's abandonment was an expression of his deepest impulses.[298] Had he chosen to remain with his family he would have fulfilled his duty at the cost of living a lie. Once again, the virtues of truth ambiguously relate to authenticity. There is a tension between the life of an authentic individual and the ethical reassurance that comes from acting out of duty rather than desire.

Williams turns to Rousseau to demonstrate why authenticity needs the virtue of sincerity. For Rousseau, self-expression is fickle and revelatory of conflicting and transitory inners states.[299] In Rousseau, Williams perceives a tension between the individual and society. Rousseau believes sincerity is at the heart of virtue, since frank speech reveals a trusting openness between two or more people, which Williams calls "the substance of living together."[300] But Rousseau's authentic expressions are frequently misunderstood. In perceiving these misunderstandings as malicious, Rousseau retreats to solitude, away from the challenges of ethical life.[301]

History also reveals what we desire from trustworthiness, which stabilizes accuracy and sincerity. I want to trust not simply that you say what you mean but that what you mean will not

---

[296] Williams, *Truth and Truthfulness*, 205.

[297] In her chapter on hypocrisy in *Ordinary Vices*, Shklar notes that in liberal societies we may prefer less frank-speech in public domains. Shklar writes: "It is, far more significantly, no longer acceptable in the United States to make racist and anti-Semitic remarks in public; yet in private conversation racism and anti-Semitism are expressed freely and frequently. Southerners use to sneer at this display of Northern hypocrisy…Would any egalitarian prefer more public frankness? Should our public conduct really mirror our private, inner selves?" (78).

[298] Bernard Williams, "Moral Luck," in *Moral Luck,* (Cambridge University Press, 1982), 20-39.

[299] Williams, *Truth and Truthfulness*, 178.

[300] Williams, *Truth and Truthfulness*, 179.

[301] Williams, *Truth and Truthfulness*, 181.

change with the wind. I want to not only expect *that* you will think and act a certain way but expect it *of* you.[302] Williams turns to Diderot, a historical figure who articulates a more sincere ideal of self-expression. In Diderot, we turn to other people to hold us accountable to our spontaneous declarations, to keep us committed to our wishes, desires, and beliefs. When I declare my beliefs, the practice of declaring them *to you* constitutes a practice that "firms up" my immediate declarations into more permanent actions in the future.[303] I recognize that you rely on my dispositions, and I want to be seen as someone you can rely on just as I want to rely on you.[304] Some people are less fragmented and more stolid in their declarations. Notwithstanding, we all require others to "steady the mind."[305]

But even Diderot's more sociable authenticity is missing something important, namely, the virtue of accuracy. Depending on whom we trust to steady our beliefs, we can easily mislead ourselves down the wrong path: "in helping you decide, I may reinforce your fantasy, and we may conspire in projecting wishes into a deceptive social hologram."[306] Authenticity therefore requires accuracy just much as it does sincerity. To act authentically, I must reflect on what is possible and on what coheres with my other commitments. Reflection, according to Williams, means asking whether my action issues from a "committed belief, supported by evidence," which means it is genuinely possible, and from a "clear-headed desire, articulated with reasons," which makes it genuinely *my* action.[307] But in reflecting on what I want and what I ought to do, I ultimately depend on other people since we rarely begin deliberation with established beliefs and desires. Practical deliberation steadies my motivations and firms up my commitments. In turn, I become clearer on which actions are authentically mine and which ones should take deliberative priority. The question for Williams therefore is not straightforwardly "What do I basically

---

[302] Margaret Urban Walker, *Moral Repair* (Cambridge University Press, 2006), 79.

[303] Williams, *Truth and Truthfulness*, 192.

[304] Williams, *Truth and Truthfulness*, 192.

[305] Williams, *Truth and Truthfulness*, 191. I wonder if the motivation to steady our beliefs is not simply a commitment to ethical relations, but a reflex against hypocrisy. As Shklar observes in *Ordinary Vices*, our aversion to hypocrisy is for many "an aesthetic necessity, a compelling need for inner purity" (46). A steady disposition curbs hypocrisy in oneself and demands the same from others.

[306] Williams, *Truth and Truthfulness*, 198.

[307] Williams, *Truth and Truthfulness*, 198.

want?" but "What do *you* think I should want?" and relatedly, "Is what I want clear-headed and grounded in reality?"[308]

Consider the richness of Williams's description of practical deliberation, particularly when compared with the State of Nature story. I note this not as a critique of the State of Nature story, but only to point out what we lose when we leave out the realistic conditions of historical genealogy. Although the individuals in the fictional state are more than just "recording devices" to the extent that they have commitments and the ability to reflect on their dispositions to truthfulness, their beliefs "embody pieces of information."[309] In real history, pooling information is a complex process of sorting out wishes from desires and desires from beliefs and is a process in which individuals rely on one another to construct and steady their motivations. Williams's genealogy aims to explain our motivations to truthfulness rather than to simply reveal the point of the practice. Our motivations are shaped in time, with others, and alongside a range of wishes, desires, and beliefs. To make sense of that account, the State of Nature story is not sufficient. We need real history.

3. Idealized Functionalist Models

Having shown why a historical sense is indispensable for Williams's genealogy, I consider two explanatory models in recent analytic philosophy, which are also interpretations of Williams's method. The first is Miranda Fricker's paradigm-based approach. The second is Matthieu Queloz's pragmatic genealogy. These interpretations, I argue, treat history as ancillary to the explanatory framework and are unified by five core features concerning: what they take to be explanatorily basic, how the explanation is constructed, what the explanation is meant to deliver, what kind of explanation it is taken to be, and the explanatory level at which functionality operates. I call these methods *idealized functionalist models.*

The first unifying feature of idealized functionalist models is that they are function-first explanations. Fricker describes this as identifying the "point" of a practice. For Fricker, even when practices are disunified, they contain a paradigm core. In her model, Communicative Blame serves as the core function of our everyday practice of blame, with its point being a

[308] See Williams, *Truth and Truthfulness*, 198.

[309] Williams, *Truth and Truthfulness*, 43.

second-personal communication that aligns moral reasons and moral understandings.[310] Queloz's pragmatic genealogy similarly grounds the explanation of a practice in the problem-solving (because need-satisfying) function it once served. In both methods, the explanations account for the *why* of a practice (i.e., its function and its emergence) rather than the *what* (i.e., its essential and accidental features).

Second, just as the function of a practice is the object of the explanation, both models employ an idealized method that constructs simplified explanations rather than complex ones. To this point, Queloz writes, "Idealization is of value here already because considering simplified prototypes of our practices in an uncluttered state of nature helps us break through the veil of familiarity to gain a sense of the more generic human purposes they serve. Idealization also cuts out noise, such as infelicitous conditions in which practices are temporarily prevented from serving a point."[311] What he means by clutter here is the "messiness of reality."[312] By stripping away "non-essential features" and by operating "with assumptions known to be false" we clarify our understanding of our practices "despite and *even because*" of these abstractions.[313]

Third, within idealized functionalist models, the explanation justifies a practice as much as it explains it. Both Fricker and Queloz draw from Williams's claim in *Truth and Truthfulness* that explanations cannot be simply descriptive but vindicatory. What makes a vindicatory explanation preferable to a descriptive one, the argument goes, is that it abrogates the ethical or epistemic anxiety that reflection on our practices stirs up.[314] This anxiety leads to a tension in idealized models: they model our practices as firmly grounded in practical reason without relying on metaphysical foundationalism or essentializing explanations of practices. Tracing normativity to functional success helps resolve this tension, or so they argue. In a reflection on Williams's genealogy, Fricker writes:

---

[310] Miranda Fricker, "What Is the Point of Blame?: A Paradigm Based Explanation," *Nous* 50, no. 1 (2016), 177.

[311] Matthieu Queloz, *The Practical Origins of Ideas* (Oxford University Press, 2021), 50.

[312] Queloz, *The Practical Origins of Ideas*, 225.

[313] Queloz, *The Practical Origins of* Ideas, 50.

[314] On the anxiety induced by reflection, see Amia Srinivasan, "Genealogy, Epistemology, and Worldmaking," *Proceedings of the Aristotelian Society* 119, no. 2 (2019), 127-156; Queloz, *The Practical Origins of Ideas*, 58; and Bernard Williams, "Philosophy as a Humanistic Discipline," in *Philosophy as a Humanistic Discipline*, ed. A.W. Moore (Princeton University Press, 2006), 180-199.

> One of the great virtues of the State of Nature method is precisely its separation of features of a concept that bear the necessity of the origin, from features that are more or less contingent matters of history. And it is entirely clear that the claim of necessity is not grounded in anything metaphysical, but, rather on something fundamentally practical—the *practical human necessity* of the materials used to construct the State of Nature scenario.[315]

Idealized functionalist models therefore explain a practice, which was not previously seen as functional, in terms of its functional role.[316] This method, as Fricker explicitly expresses, provides the reassurance of practical necessity while also acknowledging a practice's heterogeneity in ethical life.

Fourth, idealized functionalist models take a naturalistic approach to practices, an approach borrowed, again, from Williams. Simply put, ethical naturalism considers how ethical life relates to the rest of human nature rather than, say, proffering a theological explanation of an ethical demand. Blame, for instance, stems from reactive attitudes, which Fricker claims are humanly basic.[317] The extant practices of blame we vindicate are the ones we link to the paradigm core, which serves the function of second-personal moral alignment, and which is continuous with our humanly basic reactive attitudes. Queloz situates his naturalism in general conditions and root needs common to all humans.[318] As the idealized conditions and needs interact, they generate a problem to which a practice responds. Once the method establishes our paradigm practice or basic root needs, the methods demonstrate the relation the function has to our extant practices.

Fifth, for both Fricker and Queloz, functionality operates on three levels.[319] First, at the social level, insofar as a functional practice satisfies a collective need. Second, at the individual level, insofar as functionality motivates individuals to adopt the practice. And third, at the explanatory level, insofar as we make sense of why agents adopt a practice by reference to functionality. The practice of raising one's hand to speak in class exemplifies a transparent functional practice: it solves the coordination problem of determining who speaks and when; its function motivates individuals to raise their hands; and we explain why the class adopts the

---

[315] Miranda Fricker, "Scepticism and The Genealogy of Knowledge: Situating Epistemology in Time," *Philosophical Papers* 37, no. 1 (2008), 48.

[316] Williams, *Truth and Truthfulness*, 34.

[317] Miranda Fricker, "What Is the Point of Blame?" 180.

[318] Matthieu Queloz, *The Practical Origins of Ideas*, 12.

[319] Queloz, *The Practical Origins of Ideas*, 55.

practice by reference to this function. The practice of blame also serves a clear function: it solves the problem of moral misalignment, its function motivates individuals to engage in practices of blame, and we explain why a group or individuals adopt a practice by reference to the point of Communicative Blame.

From here, Fricker and Queloz's methods begin to diverge. In Fricker's paradigm model, an extant practice always serves a core function. It always has a point. Accordingly, Fricker's model first establishes a paradigm form of a practice that is "spontaneously displayed in the most simple and natural kind of blame-like exchange."[320] She then demonstrates how various practices "naturally grow" from the core function.[321] First-personal blame, self-blame, is a clear reflective instance of the second-personal paradigm of Communicative Blame. Other, less obvious forms of blame also contain this communicative core, such as in third-personal blame.[322] Even in forms where we dispassionately blame a third party from a distance there still remains "a significant residue of the basic communicative purpose" and as such "a derivative relation to our paradigm."[323] If, on the other hand, we cannot locate the point of a practice in its paradigm form, we would be justified in releasing ourselves from the moralizing habits of judgement placed on us, which are tied to contingent styles of blame rather than what is humanly necessary.[324]

Fricker argues that extant practices develop out of a paradigm. But this can only be a metaphor for explanatory priority.[325] Nowhere does she situate the paradigm of Communicative Blame on a time-axis. It simply stands as an expository device for isolating functional features already operative in extant practices. Because this approach does not justify a practice based on its origin, it avoids the genetic fallacy, or the confusion of genesis and justification. As such, Fricker argues that a paradigm-based model is preferable to a vindicatory genealogy since it

[320] Miranda Fricker, "What is the Point of Blame?," 177.

[321] Fricker, "What Is the Point of Blame?" 177.

[322] Fricker, "What Is the Point of Blame?" 178.

[323] Fricker, "What Is the Point of Blame?" 178.

[324] Fricker, "What Is the Point of Blame?" 180.

[325] Fricker, "What Is the Point of Blame?" 177.

achieves the same "explanatory payoff" more directly by relating paradigm practices to extant ones while also deflecting the charge of the genetic fallacy.[326]

Fricker's paradigm model is primarily focused on conceptually clarifying the point of our extant practices rather than on achieving something like empirical adequacy, particularly with respect to their emergence and development. The justificatory force of Fricker's paradigm therefore rests on the conceptual relation between paradigm practices and our contingent ones. When extant practices do not conceptually relate to the paradigm we are justified in recusing ourselves from their hold over us. Queloz, by contrast, argues that genealogy needs an empirically grounded explanation of how our local needs grow out of more universal ones. Whereas Fricker's models are static and ahistorical, Queloz claims we ought to model practices as dynamic and historical.[327] For this reason, I set Fricker aside for the purposes of my argument and turn my focus on Queloz's idealized functionalist model.

Within this interpretation, a model begins with a strongly idealized scenario in which a group has generic needs that generate a problem the practice then solves.[328] The model then draws on "historical information" in order to move from an idealized "proto-practice" to a de-idealized "target practice."[329] Because a paradigm model cannot explain practices that lack obvious or demonstrable function, when the paradigmatic core of a practice is lost, so too is our ability to explain it in functional terms. Queloz's model makes up for this shortcoming by hypothesizing an idealized scenario and then demonstrating how its generic conditions have undergone "complex historical elaboration."[330] Queloz's pragmatic genealogy therefore makes sense of the instrumental and non-instrumental elements of a practice as a relationship of *self-effacing functionality.*[331] In self-effacing functional practices, it is a structural feature of the practice that its function as a normative force fades into the background. Queloz differentiates

---

[326] Miranda Fricker, "Forgiveness—An Ordered Pluralism," *Australasian Philosophical Review* 3, no. 3 (2019), 243.
[327] Queloz, *The Practical Origins of Ideas*, 12; 54.

[328] Queloz, "Revealing Social Functions Through Pragmatic Genealogies," in *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*, ed. Rebekka Hufendiek, Daniel James, and Raphael van Riel (Routledge Press), 202.

[329] Queloz, *The Practical Origins of Ideas*, 51; Matthieu Queloz, "How Genealogies Can Affect the Space of Reasons," *Synthese* 192 (2020), 2008.

[330] Queloz, *The Practical Origins of Ideas*, 63.

[331] Queloz, *The Practical Origins of Ideas*, 54.

self-effacing functional practices from the contingently effaced and the necessarily effaced. In contingently effaced practices, functionality persists only accidentally, such that making the function explicit leaves present motivations intact. When we make necessarily effaced practices explicit, we erode the very motivations that sustain their expression. Making self-effaced functional practices explicit, by contrast, generates "reasonable confidence," rather than "mere confidence" by reinforcing "our sense of the practice's necessity."[332] In other words, the normative force of a practice is strengthened on all three levels (collective, individual, explanatory) only after we perceive the practice's point. What is more, Queloz argues that modeling these dynamics reconciles a pure functionalist view, in which we explain practices solely in terms of needs, with an anti-functionalist view, in which we altogether purge functionality from the normative and explanatory force of practices.[333]

Pragmatic genealogy, so conceived by Queloz, explains "both why we have an idea at all, in some form or other, and why we have it in the specific form it now has around here."[334] The point of genealogy, then, is to identify and relate the generic and local needs to which our practices respond.[335] Crucially, only pragmatic genealogy can accomplish this task. As Queloz notes, "If one were specifically and *exclusively* interested in whether a practice serves some particular and highly local need, one would not need genealogy."[336] In such cases, a paradigm model would suffice. The explanatory priority of pragmatic genealogy—and it seems, of genealogy in general—is to reveal the practical necessity of our extant practices, rather than to de-sediment the historical deposits of history layered upon them.

Idealized functionalist models describe the function of practices in simplified models, and appeal to ethical naturalism so as to distinguish between universal and contingent human needs and practices. Does history play a role here? For Fricker, making sense of local, contingent practices does not require history. A realistic paradigm model is explanatorily sufficient.[337]

---

[332] Queloz, *The Practical Origins of Ideas*, 56.

[333] Queloz, *The Practical Origins of Ideas*, 54-55.

[334] Queloz, *The Practical Origins of Ideas*, 68.

[335] Queloz, *The Practical Origins of Ideas*, 63.

[336] Queloz, *The Practical Origins of Ideas*, 63.

[337] Fricker, "What is the Point of Blame?," 165.

Queloz's model draws on "historical information," though does not incorporate historiographical methods, that is, actual empirical inquiry about history.[338] For both Fricker and Queloz, empirical history is at best absent and at worst subordinated to idealized functionalist models.

To make matters worse, Queloz downplays the role of history not simply in pragmatic genealogy but also in Williams's own genealogical method, methods he views as one and the same. In Queloz's interpretation, Williams's genealogy is "imaginary, vindicatory, and revelatory of practical necessity" before it is "historical, critical, and revelatory of contingency."[339] When Williams is engaged in history, it is no more than a "breezy romp," and fails to offer "anything like the historian's craft."[340] It is one thing to fault Williams for neglecting historical detail, a charge that admittedly carries some weight. It is quite another for Queloz to dimmish the importance of history in his own method.

Undoubtedly, Queloz provides a lucid and insightful reading of Williams. However, the irreverence he pays to historiography in *Truth and Truthfulness*, plunders genealogy of its humanistic spirit and ignores Williams charge to philosophers in "Philosophy as a Humanistic Discipline":

> [P]hilosophers cannot altogether ignore history if they are going to understand our ethical concepts at all. One reason for this is that in many cases the content of our concepts is a contingent historical phenomenon… [T]he reflective understanding of our ideas and motivations, which I take to be by general agreement a philosophical aim, is going to involve historical understanding. Here history helps philosophical understanding, or is part of it. Philosophy has to learn the lesson that conceptual description (or, more specifically, analysis) is not self-sufficient; and that such projects as deriving our concepts *a priori* from universal conditions of human life though they indeed have a place (a great place in some areas of philosophy than others), are likely to leave unexplained many features that provoke philosophical enquiry.[341]

In much of his work Williams cautioned against contemporary philosophy's tendency to "get things right" using the techniques of the natural sciences rather than the humanities.[342] Though

---

[338] Queloz, "How Genealogies Can Affect the Space of Reasons," 2008.

[339] Queloz, *The Practical Origins of Ideas*, 188.

[340] Queloz, *The Practical Origins of Ideas*, 10; 172.

[341] Bernard Williams, "Philosophy as a Humanistic Discipline" in *Philosophy as a Humanistic Discipline*, ed. A.W. Moore (Princeton University Press, 2006), 191-192.

[342] Nicholas Smyth, "Purity and Practical Reason: On Pragmatic Genealogy," *Ergo: An Open Access Journal of Philosophy* 10, no. 37 (2023), 1057-1081.

sensible for some topics, the sciences serve as a bad model for ethics.[343] I refer to contemporary philosophy because, for Williams, figures in the history of philosophy—Plato, Hobbes, Hume, Rousseau, and Mill are some of his examples—often paired the force of reason in their arguments with "imaginative and expressive powers."[344] For them, as it should be with us, conversion stems not simply from the argumentation being true, but its "ringing true."[345] "The demand" Williams writes, "that moral and political philosophy should sound right, should speak in a real voice, is not something arbitrarily imposed by those with a taste for literature, or for history, or for excitement. It follows from philosophy's ideal of reflectiveness, an ideal acknowledged in the subject's most central traditions."[346]

Philosophical argumentation is the art of "truthful style."[347] For this reason, among many others, history and philosophy are natural compatriots. They both reflect on and depict the life of society. They attempt to make sense of individual and collective motivations and do so in a way that involves discovery and a passion for getting it right. History and philosophy are also interpretive disciplines and interpretation involves style. In this regard, "the force of reason" and "the power of persuasion" cannot be separated.[348] In genealogy, history and philosophy also complement one another in their respective limitations. Philosophy offers a critical framework for reflecting on and clarifying our values and practices, while history provides the temporal framework for understanding how those practices have developed.[349] What is more, history offers a descriptive account of motivations and practices, which philosophy, in turn, can normatively orient.

---

[343] See Bernard Williams, "What Might Philosophy Become," in *Philosophy as a Humanistic Discipline*, ed. A.W. Moore (Princeton University Press, 2996), 200-213; Williams "On Hating and Despising Philosophy", *London Review of Books*, April 18, 1996. See Smyth, "Purity and Practical Reason," for his remarks on why borrowing methods from the natural sciences does not work.

[344] Williams, "What Might Philosophy Become," 205.

[345] Williams, "What Might Philosophy Become," 206.

[346] Williams, "What Might Philosophy Become," 205-206

[347] Bernard Williams "On Hating and Despising Philosophy," *London Review of Books*, April 18, 1996.
[348] Williams, *Truth and Truthfulness*, 226.

[349] See Colin Koopman, "Bernard Williams on Philosophy's Need for History," *The Review of Metaphysics* 64 (2010), 3-30.

In his commentary on Williams, Mark Jenkins observes that no other intellectual figure "ancient or modern" has received as much praise from Williams as Thucydides has, particularly in *Shame and Necessity* and *Truth and Truthfulness*.[350] What Williams draws from Thucydides is the idea that explanations of events are not saturated with ethical or moral reasons. The interpretive framework he provides is realistic and naturalistic insofar as he explains human motivation in its entirety. Considered thus, Williams notes an impartiality in Thucydides. Impartiality here is meant not as a value-free perspective, but as an effort to make sense of social events by "relating them intelligibly to human motivations, and to the ways in which situations appear to agents."[351] It is hardly coincidental that the figure from whom Williams draws so extensively to make sense of human motivation is a historian. This is because, for Williams, philosophy cannot make sense of the full range of human motivations, or at least not without a more humanistic interpretive framework to aid its inquiry.

4. Realistic Explanations

With a clear account of both vindicatory genealogy and idealized functionalist models in view, I now consider which of these methods is best suited to the aims of a politically realist inquiry. At first pass, Williams's genealogy is the clear choice. Nearly all realists emphasize the importance of history for political philosophy, so an explanation that involves itself with history is obviously preferred to one that does not. What is less obvious, I argue, is why history provides the insights we need for a politically realist inquiry. To answer this, I introduce the notion of empirical adequacy, which allows us to adjudicate between competing explanatory styles and to show why historical genealogies are better suited to the aims of political realism.

An explanation is an account that describes or makes intelligible the why or how of a phenomenon. In the context of practical deliberation, an explanation is the reason or motivation given for an action. An explanation also aims to depict the phenomenon as accurately as possible. It aims, in other words, to capture a reality. When we speak of things "as they really are," the implication is that there is an order not subject to the will. Williams describes this basic sense of reality in terms of "obstance," though he notes that in many cases obstance is so strong

[350] Mark P. Jenkins, *Bernard Williams* (McGill-Queen's University Press, 2006), 165.

[351] Bernard Williams, *Shame and Necessity* (University of California Press, 1993), 161.

that an agent might struggle to imagine how she could even exert her will.[352] Meeting someone on the moon in thirty seconds time or catching yesterday's train are some examples he provides.[353] Try as I might, there is something independent of my will that prevents me from turning my motivations into action. Indeed, something prevents me from conceiving of how I would even attempt to meet you on the moon or catch yesterday's train. It is this general connection between the will and resistance to its determination that delineates what Williams means by reality here.

It is one thing to concede the existence of metaphysical or logical realities independent of the will, or at least in cases where our intuitions easily converge, as I suspect they likely will in the examples above. It is another matter to consider what, if anything, constitutes a political reality and its corresponding forms of resistance or obstance. And yet, there are phenomena in politics that resist the will. Suppose, for example, someone demands her state representative eliminate government corruption. If the politician replies that such demands are impossible, her constituents might accuse her of giving up and capitulating to the status quo. But is the accusation based in political reality? After all, what would it look like to try here? Overhauling bureaucratic systems? Eradicating entrenched incentives to lie, cheat, or exploit procedural gaps? Cultivating an airtight civic ethos among elected officials? Setting that aside, one might plausibly outline steps she could take to mitigate corruption, particularly on a localized scale, but the broader demand to eliminate it *tout court* will meet real resistance. Obstance here is not the same as they are in Williams's examples, but the present conditions are such that no individual action could bring about the desired outcome. The demand, as it stands, is not realistic.

Imagining political impossibilities is not inherently problematic. Utopias expand our imagination and can prove quite helpful when comparing two states of affairs, a topic I discussed in Chapter Two. The danger of following these impossibilities arises when we confuse their fictional status with reality.[354] Williams calls this tendency for fiction to subvert true belief *wishful thinking*, which consists of individuals (or groups) forming their motivations on the basis of what might be pleasing to imagine rather than on what is actually possible. In the contexts of

[352] Williams, *Truth and Truthfulness*, 136.

[353] Williams, *Truth and Truthfulness*, 136.

[354] Williams, *Truth and* Truthfulness, 139.

politics, wishful thinking reflects a conflict between our wish for a radically improved world and our true belief about that world.[355]

Not all wishes subvert true belief or serve as what Williams calls a "surrogate of reality."[356] This relates back to the idea of obstance. Williams considers an example of basic arithmetic: suppose someone is required to pay a bill for $3,000 but only has $2,500 to his name, with no access to credit or overdraft. He might wish that $2,500 minus $3,000 were not a negative number though he could not reasonably conceive of this as a real possibility.[357] To preserve his sanity, he must consciously recognize his wish as a wish. However, the wish can take hold of reality by anchoring itself in some part of reality. The original wish that $2,500 minus $3,000 is not a negative number is now replaced by a more plausible surrogate. For example, he might convince himself that he will make more money before the check clears by gambling or through some other get-rich-quick scheme. According to Williams, the surrogate retains some of the original wish's features but sheds the stronger impossibility. Eventually, however, reality resists the fantasy as it inevitably must.

In the case of practical impossibilities, the structure of belief formation is the same. Consider the reality that agents tend to fall short of ideal moral reasoning. Generally speaking, agents often weigh their wishes, desires, and prudential concerns alongside moral ones. Now, an idealist might face this reality and wish agents gave moral considerations absolute priority in deliberation. Anyone who shares her values will sympathize with this wish. However, an idealist who is honest with herself, or honest enough, will recognize the wish as a fantasy. The problem, therefore, is not the wish itself, but the way it takes over reality by adopting both elements of fact and fantasy and how it alters our expectations of agential action. For example, the philosopher may now come to claim that, with the right institutional support, agents will reason and act morally. This modified wish is more conceivable because we do observe such behavior in

[355] We should not confuse wishful thinking with magical thinking, which is the psychological disposition to causally link two unrelated events together, despite no evidence of any plausible link between them (e.g., if I imagine my dissertation complete, it will write itself). Although both wishful and magical thinking involve belief formation that subverts reality, the former concerns excessive hope subverting true belief, whereas the latter refers to a belief's ability to causally determine reality. When one falls victim to wishful thinking, she does not simply believe her hope and optimism will make reality thus and so, but believes her optimistic view of the world *is* how things really are, or really can be, despite evidence indicating otherwise.

[356] Williams, *Truth and Truthfulness*, 137.

[357] Williams, *Truth and Truthfulness*, 138.

localized scenarios. As the surrogate becomes more plausible, it makes room for what Williams calls *effective fantasy*.[358] Once effective, the fantasy obscures its own status as fantasy and begins to function as if it were reality, thereby distorting the formation of true belief. In this way, the fantasy does not merely coexist with reality but *becomes* reality. As it is typical of self-deception, these wishful thinkers do not blame their sense of reality when the evidence supporting their view fails to materialize. Instead, they blame those who fail to exercise their moral reasoning or who give up on these principles and succumb to the status quo.

Wishful thinking can just as easily work in the opposite direction, towards dismay. The cynic might come to believe that human beings are, by nature, incapable of ethical conduct. As with the idealist's impossible wish, engaging with the world will eventually compel the cynic to acknowledge that people do in fact act ethically. In response, she may construct a surrogate of the impossible, such as the claim that, when faced with a choice between prudential and ethical considerations, an agent will reliably choose the prudential route. We can see how this wish can take hold of reality and become effective fantasy. Since the surrogate quite often appears more politically and psychologically plausible it allows the cynic to maintain her outlook despite counterevidence.

Strange as it may sound, I do not believe it is misplaced to describe this kind of cynical outlook as wishful thinking. Those who see little hope in the possibility of ethical life may, in some cases, wish it were so, not because they take any straightforward pleasure in cynicism. Being right, even about unfortunate states of affairs, offers its own kind of consolation. Rather than revise their impossible view in light of evidence, both idealists and cynics construct a surrogate that preserves the fantasy while appearing more realistic. There is a certain pleasure one gains in maintaining a fantasy, whether to avoid the pain of being wrong, or because reality is just too difficult to bear.

Effective fantasy subverts the formation of true belief. It does not yield effective ideals, since its prescriptions remain rooted in wishful thinking. For Williams, resisting such thinking requires cultivating skills and attitudes that counter this satisfaction.[359] The question then, is *how* does one do so? What dispositions and values enable an inquirer to arrive at true belief, such that

[358] Williams, *Truth and Truthfulness*, 139.

[359] Williams, *Truth and Truthfulness*, 125.

she gains a sense of reality? When choosing between two explanations or two methods, what kind of criteria should an inquirer adopt to distinguish between the two? I argue that we ought to draw on the idea empirical adequacy understood as the degree to which an explanation matches observed phenomena. Although we should not expect an empirically adequate explanation to offer a comprehensive account how things really are, we still can develop a method that, when developed with care and attention, produces a realistic explanation. An empirically adequate explanation does not require one definitive method, though there are better and worse inquiries. When choosing between competing genealogical methods, I propose four general criteria we should draw on in adjudicating empirical adequacy: determinacy, impartiality, contingency, and regularity.

First, the explanation must place the conditions that structure human motivation and action in a *determinate* time and place.[360] Since the explanation intends to say something about our practices rather than fictional practices, the account must relate motivations to determinate social events and structures. It is not enough to assume or posit that the explanation approximates real conditions or that conditions could obtain in real life.[361]

Second, an explanation should be *impartial* in the sense that what the audience prefers to hear or already agrees to should not place undue pressure on how one selects and constructs the material of the explanation itself. Williams contrasts an accurate explanation to a mythical or fictional story, which crafts its narrative according to what would best suit the audience.[362] An accurate explanation takes no such considerations into accounts. If the explanation adds or omits details to please the audience or because it wants to find some common ground of agreement, then its accuracy should be called into question. Although the explanation may be aware of its audience, it should not tailor the contents of its account to suit the audience's desires or beliefs.

Third, the explanation must account for a large degree of *contingency*. As Williams notes, an easy way of misunderstanding practices and, by extension, the motivations expressed in those practices, is to over-rationalize them.[363] He writes, "many historians do seem to overestimate the

[360] See Williams, *Truth and Truthfulness*, 162.

[361] Nicholas Smyth, "Purity and Practical Reason: On Pragmatic Genealogy," 1065.

[362] Williams, *Truth and Truthfulness*, 165.

[363] Williams, *Truth and Truthfulness*, 248.

extent to which people knew what they were doing," which means they tend to reconstruct narratives that make past practices seem "more sensible than they actually were."[364] The same goes for philosophers who, in making sense of practical deliberation, characterize human motivations as far more rational than they actually are. Tempting as it may be to present a simplified account of practical deliberation, an explanation without contingency easily becomes reductive. There are several problems with reductive explanations, one of which is that they risk over-rationalizing past action by suggesting that those who acted otherwise did so merely because they were wrong or irrational. Reductive explanations will also fail to capture the ambiguity in how situations appear to agents during practical deliberation.

Fourth, because an explanation accounts for the process of practical deliberation, we must understand that process as coherent rather than as a string of haphazard events. This requires a *principle of regularity*, that is, a mechanism, to explain how motivations become expressed in practice in a way that is not simply contingent.

Although I argue that all four criteria are required for a realistic genealogical explanation, I leave room for disagreement over their relative weight. Even where genealogists agree on these criteria as requirements of adequacy, they may differ over how much importance to assign to each. Some approaches may place greater weight on contingency, while attaching less importance on to the articulation of regular patterns or processes governing motivation and practice, or vice versa. Having defined the criteria for a realistic explanation, I now consider idealized functionalist model, and more specifically Queloz's model against Williams's genealogy. Although Queloz has the ambition to deliver an empirical explanation of our practices, his model does not satisfy the four criteria for empirical adequacy.

5. Explanatory Choices

Idealized functionalist models blend philosophical explanation with justification. As hybrid accounts, idealized functionalist models occupy an intermediate space between causes and reasons insofar as they purport to justify particular practices by showing that those practices are rational because they satisfy basic human needs. This hybrid structure gives rise to two related objections to genealogies understood as explanatory: the genetic fallacy objection and the continuity objection.

[364] Williams, *Truth and Truthfulness*, 248.

When one commits the genetic fallacy, she has conflated the origin of a practice with its justification.[365] Against this critique, Queloz argues "we need to question the traditional role allocation according to which the state of nature merely justifies while real history merely explains. We need to make room for the idea that pragmatic genealogies can be genuinely explanatory despite operating with a fiction, and that they can be normatively significant *despite* being explanatory."[366] This raises then need to assess how Queloz refutes the genetic fallacy and whether his refutation raises further concerns.

If an idealized functionalist model only aimed to yield a justification, then it would not be so encumbered with reconciling the fictional and the real, the hypothetical and the actual. Queloz's method could resolve this tension by remaining idealized from beginning to end. In this sense, the justification could still draw on contingent values, but it would not entangle itself in the genetic fallacy. This is a perfectly acceptable justificatory model since it helps us understand our practices in relation to rival ones.[367] A justification of this kind would be similar to the non-concessive idealizations I discussed in Chapter Two. From a non-concessive view, chimerical assumptions clarify our intuitions and tests the relationship between competing states of affairs. In such a justificatory model we would measure one practice against another or against a world in which the practice does not obtain. Whether these worlds are practically or deeply impossible does not matter so long it aims to justify and not explain. But this is not the case. Indeed, Queloz's model requires an explanatory feature because, as I mentioned, the reasons we have to endorse a particular practice derives from an explanation of our practical reasons to adopt it in the first place.

Queloz locates the normative force of the idealized functionalist model in the formation process itself rather than in the origin of the practice.[368] In other words, the formation process of a practice is itself reason giving. Queloz writes:

> We can acknowledge that there is such an error without committing ourselves to the much stronger claim that nothing can be inferred about the justification of something from facts about its origins. Items in the context of formation can form part of the context of justification, but they can properly do so only if there

[365] See Colin Koopman, *Genealogy as Critique: Foucault and the Problems of Modernity* (Indiana University Press, 2013), 62-65.

[366] Queloz, *The Practical Origins of Ideas*, 12.

[367] Queloz, *The Practical Origins of Ideas*, 220.

[368] Queloz, *The Practical Origins of Ideas*, 121.

> is a connection between some aspect of the context of formation and the justification of the item in question…What normatively ambitious genealogical explanations depend on, then, is that there be a connection rendering some aspect of the context of formation relevant to the context of justification.[369]

Queloz also notes that what we see as counting as a reason for or against a practice is contingent on a whole range of influences, particularly that of our first-order commitments.[370] Our standards come from the inside, so what we accept as a reason for action will depend on what we contingently value.

If we grant that Queloz's model refutes the genetic fallacy objection in the manner described above, it faces the *continuity objection* as raised by Nicholas Smyth. The continuity objection concerns the inference a model can draw between the explanatory features of a model, which relates to the hypothetical state of nature, and the justificatory features of a model, which relates to our present practices.[371] According to Smyth, if a model demonstrates that a practice emerged to solve a practical problem in some hypothetical state, it does not therefore suffice to show that it now has that same point and, subsequently, that we now have the same reason to endorse it. The only scenario in which that inference can be drawn is when we can demonstrate that the conditions of the hypothetical state are continuous with our real conditions, such that the point is responding to the same practical pressure.[372] In other words, to justify the idealized practices as practices we ought to endorse, the model must show how continuity between the idealized account and our actual present obtains. But as Smyth notes, we cannot establish this continuity on intuition alone: "we have to *check* whether suitably specified descriptions match, and this involves doing a great deal of empirical inquiry into our own society, into the ways it matches (or fails to match) the precise descriptions we needed to give of the ideal world."[373] Considering the idealized model will have to draw on empirical methods at some point, Smyth wonders whether we can circumvent this step (and many of its problems) by dispensing with idealization altogether and instead beginning with empirical genealogy.

---

[369] Queloz, *The Practical Origins of Ideas*, 214.

[370] Queloz, *The Practical Origins of Ideas*, 216.

[371] Smyth, "Purity and Practical Reason: On Pragmatic Genealogy," 2023.

[372] Smyth, "Purity and Practical Reason: On Pragmatic Genealogy," 1057.

[373] Smyth, "Purity and Practical Reason: On Pragmatic Genealogy," 1065.

Smyth's argument leads to a deeper difficulty, which is related to the determinacy criterion for empirical adequacy. Idealized functionalist models are, at least in part, indeterminate, which entails we have no means of drawing a continuity between the hypothetical needs and our present conditions.[374] Queloz writes "We can organize and elucidate the complex amalgam that is our target practice using a simplified proto-practice. And we can then also model how we might have gotten from here to there, *where 'there' does not refer to some datable moment of emergence*, but to an abstractly characterized basic predicament of which our present situation is a socio-historically local manifestation."[375] Although Queloz's point is that continuity can be constructed between a proto-practice and our present situation, it remains unclear how such a comparison is meant to yield sufficient explanatory or justificatory force.

Queloz suggests that his model can avoid continuity failures by operating at a high level of abstraction and by anchoring present practices in basic human needs.[376] Yet this strategy introduces another problem for both the explanatory and justificatory ambitions of the model. It remains unclear how one's recognition of the practical point of a practice within an idealized framework will adequately motivate her to adopt that practice under realistic conditions. Absent the concrete context in which a practice and its associated function make sense, it is difficult to see how we could take up or sustain a practice. Genealogy thus formulated would be non-concessive in the utopian sense. This difficultly parallels familiar problems in Rawls's justificatory model, though what saves Rawls's original position from the non-concessive defect is it explicitly justifies the principles of justice and does not explain patterns of human motivation or conduct. Queloz's idealized model thus faces challenges twice over insofar as he aims to justify practices in terms of his explanation.

One way Queloz tries to address this dual problem is by establishing continuity through a staged progression. He first begins with general structural needs and then piecemeal introduces localized needs. This method demonstrates how one practical exigency gives rise to the next, eventually culminating in our present practice. At this point, however, the model begins to lose

[374] Williams's account may not be as vulnerable to the continuity objection since he repeatedly insists that a vindicatory genealogy operates with distinct kinds of explanation—the fictional and the real—and takes care to keep the two separated.

[375] Queloz, *The Practical Origins of Ideas*, 63; emphasis mine.

[376] Queloz, *The Practical Origins of Ideas*, 228.

any genuine historical sense since Queloz treats the relation between the hypothetical proto-practice as one of logical entailment, such that only those historically specified practices that are logically entailed by earlier stages are built into the account. What is more, Queloz maintains that we can work out the logical entailment between the structural and contingent needs and practices, without worrying about whether the inferences we draw are strictly speaking empirically adequate. If it does turn out the inference between the generic and local are wrong, Queloz writes, "the problem will then not be that it has subtly trespassed against the canons of reasoning, but simply that it is unsound."[377] *Simply* unsound? Granted, fallibilism requires the genealogist recognize her interpretations may be mistaken or that she may have missed important details, a regrettable but inevitable part of inquiry. Yet, openly conceding that the basic premises of one's explanation may be empirically false while insisting that this in no way undermines the method, misses the point of genealogy. Smyth's continuity objection therefore brings into view idealized functionalist model's failure to explain how the conditions, needs, and practices posited by the model determinately relate to those of the present. At best, the model can establish relations of logical entailment. It cannot trace the empirical process of development between the two. This in turn brings into focus a deficit of contingency in accounts of the emergence and development of our extant practices.

A further difficulty arises, moreover, with respect to how Queloz disambiguates the explanatory aims of his model from its justificatory aims. On the explanatory side, Queloz's model seeks to demonstrate why a practice exists by identifying the practical problem it helps solve and how that practice has developed from solving structural problems to more contingent ones. On the justificatory side, Queloz's model aims to affect the space of reason through logical inference from the generic structural needs or our local, contingent ones. Queloz's model claims to be explanatorily powerful because the idealized proto-practices and structural needs are universal, and "inscribed in human nature." [378] They point to something deep in human nature that goes beyond more subjective wishes, desires, and preferences, all of which are shaped by contingent factors. If I had been raised another way, I would not have my preferences or desires,

[377] Queloz, *The Practical Origins of Ideas*, 231.

[378] Queloz, *The Practical Origins of Ideas*, 229.

but I would certainly have the same needs.[379] In this regard, the explanatory elements of his model are impartial.

This is not the case, however, with his justificatory model. Consider Queloz's insistence that any genealogy must frame its inquiry in terms of its audience, by asking "to *whom* it is addressed. Who needs to hear it? Whom does it aim to convince."[380] He here echoes Williams own description of justificatory structures in *Morality* and *Ethics and the Limits of Philosophy*.[381] Just as Williams does not seek to persuade the amoralist when offering an ethical justification and instead addresses those who already demonstrate ethical concern, so too does Queloz maintain that genealogy should not function as "an instrument of conversion."[382] Instead, genealogy promotes self-understanding by reinforcing the confidence of those already inclined to value an extant practice for its own sake. If Queloz's view here is simply that pragmatic genealogy takes its audience into consideration, in the sense that it directs its explanatory and justificatory model to a particular set of people or expects its justification to ring true with some rather than others, then I can fully accept his argument. However, I suspect that Queloz is arguing for more than the trivial truth that our audience comes to us with values in hand. So, how might we frame this relationship between genealogy and its audience and how might that affect the accuracy of the explanation itself?

Queloz's conception of an audience-relative relationship might emerge from Williams's own reflections on historical interpretation. By specifying structural needs and predicaments "any creatures like us would be likely to face" genealogy provides a minimal explanation that most people would be compelled to concede.[383] Put differently, a genealogy takes its fictional model to grow out of uncontroversial claims about anthropological facts. On this point, Williams writes, "If it is claimed that some supposed statement of fact in either a chronicle or a narrative is too interpretative or carries too many presuppositions with it… we can retreat to less contentious

---

[379] Queloz, *The Practical Origins of Ideas*, 224.

[380] Queloz, *The Practical Origins of Ideas*, 182.

[381] Bernard Williams, "The Amoralist," in *Morality: An Introduction* (Cambridge University Press, 1972), 3-13; "The Archimedean Point," in *Ethics and the Limits of Philosophy*, 22-29.

[382] Queloz, *The Practical Origins of Ideas*, 182.

[383] Queloz, *The Practical Origins of Ideas*, 63.

statements and work from there."[384] Now consider Queloz's description of the "developmental narrative of genealogy":

> In the genealogical mode, we can present these fairly sophisticated needs as growing out of more primitive needs, and these out of even more primitive ones, until we reach *what our audience is most likely to be willing to regard as uncontroversial needs*—needs that are uncontroversially the genuine article (and not merely interests masquerading as needs) and that human beings uncontroversially have anyway.[385]

Much like Williams argument that we resolve interpretive disagreements by retreating to more minimal grounds, Queloz's method builds upon needs he argues we would all agree to. If we find ourselves in disagreement over what constitutes a structural need—for instance, if some view a need as really a preference—then we can retreat to first-order needs and begin to build the model out again. Williams story, for example, begins with the uncontroversial need for information and cooperation before proceeding to derive local needs we initially did not recognize *as needs*, such as the practical value of accuracy and sincerity. If, however, someone disagrees that the cooperative pooling of information is a structural need, we would retreat to first-order needs, such as the need for safety.

It is less clear whether Williams also holds this justificatory relationship with the audience in his historical account. If so, it seems at odds with how he frames accurate explanations in general. Consider, once more, Williams's description of the mode of myth: "the question whether the story should be told is just the question whether the story is appropriately directed to its audience, whether, as I put it before, it will suit them."[386] Williams's description here resonates with the justificatory structure of Queloz's genealogy. But, as Williams argues, the mode of historical truth is not audience-relative. It does not bend its narrative to the audience's will, nor does it aim to tell a story the audience is most likely to agree to. Of course, Williams notes that different narratives might invite different interpretations. For example, feminist histories might show that the previous interpretations of the past, which we thought were objective, were in fact deeply sexist. A feminist historian will provide a new interpretation that seeks to dismantle these past inaccuracies and her interpretation will be welcomed by those who

[384] Williams, *Truth and Truthfulness*, 239.

[385] Queloz, *The Practical Origins of Ideas*, 65; emphasis mine.

[386] Williams, *Truth and Truthfulness*, 165.

share her values. Nevertheless, we can also imagine the feminist historian carrying out her inquiry not in service of her feminist audience, but in service of truth.

Idealized functionalist models draw explanation and justification too closely together. Perhaps it is the case that the explanatory dimensions of Queloz's model are audience-impartial in the sense I have described. They do not merely aim to tell a story that the audience will agree to but to capture something about reality. By contrast, the justificatory dimension may be audience-relative insofar as tailors its justification to the audience's context and prior commitments. But there is great difficulty in reconciling explanation and justification into one seamless method. Each involves distinct commitments to evidence and to its audience. When a model maintains a separation between descriptive explanations and normative justifications it avoids the genetic fallacy and provides better explanations.

Finally, Queloz's accuracy is hindered, in part, by his methodological commitment to perspicuous model-building. To the contrary, it is evident that Queloz believes that uncluttered models produce better explanations because they strip away the unnecessary or inconvenient elements of his explanation. But a simplified model showcases practical reason as we wish it would be, rather than as it actually is. Smyth describes this disciplinary ethos as a "drive for purity in practical philosophy."[387] One might suppose that Queloz favors criteria used to assess scientific theories, such as simplicity; however, a central criterion for assessing explanatory power in the sciences is empirical adequacy. Queloz's model does not meet this standard, nor does empirical adequacy seem to be a criterion his model meaningfully prioritizes.

Does Williams's vindicatory genealogy fail to meet these criteria for empirical adequacy as well? For the most part, no. Given that Queloz considers his model an interpretation of Williams's method, I suggest the crucial difference between Queloz and Williams's respective methods is that the latter views the purpose of the fictional and the real parts of genealogy as distinct in kind rather than in degree.[388] Indeed, part of Williams's critique of Nietzsche's genealogy is that he fails to disambiguate between historical fact and functionalist abstraction.[389]

---

[387] Smyth, "Purity and Practical Philosophy: On Pragmatic Genealogy," 1078.

[388] For a description of the fictional and empirical dimensions of Williams's genealogy as being a difference of kind rather than of degree, see Colin Koopman, "Genealogy, Empirical History, and Functional Teleology." *Mind*, November, 2025 https://doi.org/10.1093/mind/fzaf038; see also Williams's discussion of the "two distinct ways of talking" between myth and history in *Truth and Truthfulness* (169).

[389] Williams, *Truth and Truthfulness*, 37-38.

For Williams, the fictional State of Nature clarifies why we need truth because it serves some basic human need in a very basic sense. The fictional story thus explains what problems something like the virtues of truth would solve. At this point, the fictional story simply explains why life would be better if individuals of that group adopted the dispositions of accuracy and sincerity. The fictional story does not explain why it is rational for individuals to adopt truthful practices, so we cannot draw the inference that it would be rational for us to do so. What is more, the account of the fictional story cannot, on its own, generate normative force on the basis of us needing the virtues of truth because they are too abstract. They are not world-guiding because they are not guided by *our world*.

For this, we need history and, more precisely, empirically adequate history. We need history to *fill in* the fictional story and provide normative guidance about what the virtues of truth look like under realistic conditions. We need history to *extend* the fictional story by tracing how contingent forces have elevated what we regard as needs and by reminding us that these practices can one day disappear. We need history to *diagnose* practices that bear no internal conceptual relation to the virtues of truth yet nevertheless stand to benefit from their integration. And because Williams incorporates a historical sense into his method, he adopts the criteria appropriate to historical explanation. His empirical genealogy is determinate in the sense that it points to real historical developments and instances of the virtues of truth, for example, in his account of accuracy within practices of scientific inquiry. It is sensitive to the historical contingency of these virtues, as illustrated by his discussion of the invention of historical time itself. It is impartial insofar as his account does not adjust the needs, conditions, or practices of real history to suit a particular audience. Finally, Williams's genealogy appeals to a wide range of psychological mechanisms beyond functionality, such as shame, honor, and respect, that are nevertheless reason giving, in the sense that they help explain why agents come to adopt the virtues of truth in real history. What Williams's genealogy is *not* is simple. It explains our motivations to adopt the virtues of truth by uncovering a complex set of reasons. These reasons can only be found under realistic conditions and through detailed historical investigation.

6. Realism and Contextual Values

The chapters thus far have sought to develop the aim of a politically realist inquiry: the construction of practicable ideals that normatively guide us in the world because they are

themselves guided by the world. An inquiry with this aim requires accurate, which is to say, empirically adequate, explanations of our practices such that the observational content of the explanation of human conduct aligns with the kinds of phenomena we in fact encounter.

In this chapter, I have specified four criteria that bear on explanatory accuracy and that structure our choice between competing explanatory accounts: determinacy, impartiality, contingency, and regularity. Taken together, these criteria articulate the standards by which genealogical explanations should be assessed if they are to support the politically realist demand for empirical adequacy. As I have argued, Queloz's idealized functionalist model is explanatorily insufficient with respect to all four criteria. It provides no means of checking the observational content of its explanation against our extant practices; its emphasis on logical entailment between generic and local practices fails to capture the contingency of our practices; its explanatory features are tailored to a particular audience; and the mechanism guiding generic and local practices are reduced to practical reason and need-satisfaction. Williams approach, by contrast, satisfies the four criteria. We should therefore prefer William's method when choosing between the two approaches.

Empirical adequacy is what we might call a cognitive or epistemic value insofar as it is conducive to yielding some truth about the world. There are also, I suggest two further contextual or non-cognitive values that equally guide a politically realist inquiry.[390] The first relates to the value of *freedom*, which expresses the desire for human potentiality and flexibility in being different than how we currently are.[391] In Chapters Two and Three I discussed how this value is central to political realism of Shklar and Williams as well as to the feminist realism of Anderson and Young. We want ideals that are informed by our actual conditions and that are suitable to the kinds of people we are, but we equally want to avoid an uncritical realism that would prevent us from creating more ethical ideals. We therefore want an explanatory framework that leaves room for our world to be otherwise. Second, an explanation does not only

[390] I draw my reading between cognitive and non-cognitive values from Helen Lonigno, "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy," in *Feminism, Science, and the Philosophy of Science,* ed. L.H. Nelson and J. Nelson (Kluwer Academic Publishers, 1996), 39-58; and Longino, *Science as Social Knowledge* (Princeton University Press, 1990); as well as Elizabeth Anderson, "Human Interests, and Objectivity in Feminist Epistemology," *Philosophical Topics* 23, no. 2 (1995), 27-58; and Anderson, "Use of Value Judgements in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce," *Hypatia* 19, no.1 (2004), 1-24.

[391] For her discussion of the importance of autonomy and creativity as a value in science, see Anderson, "Human Interests, and Objectivity in Feminist Epistemology," 50.

tell us why our practices are this way, but also how they have developed over time and become realized in our conduct. Rather than explaining the "how" in simple terms, a politically realist inquiry emphasizes what I call *complexity of interaction*, which is an approach to explanatory frameworks that eschew single factor causal accounts for ones that incorporate more dynamic interaction.[392]

The criteria of empirical adequacy to which freedom and complexity of interaction apply are contingency and regularity. In light of these additional non-cognitive values, we can now assess whether and to what extent Williams's method is best suited to a politically realist explanation, particularly among rival explanations. Although Williams provides an explanatory framework that accounts for the contingency and regularity of the emergence and development of our practices, his method lacks these further contextual values. One reason Williams's genealogy fails on this front is that he draws much of his historiography from Thucydides, who frequently explains events and actions in terms of the psychological laws that undergird them.[393] A comparable appeal to timeless givens taints Williams's State of Nature genealogy: truth, the subject, and a rudimentary human nature grounded in universal human needs.[394] Even if these givens do not inhabit a robustly metaphysical framework, as they arguably do in Thucydides, they nevertheless take explanatory priority for Williams's method. What regularly compels agents to adopt the virtues of truth are deeply psychological. Shame, humility, honor, and respect are the mechanisms that produce truthful conduct. Williams pays little attention the structures, networks, or techniques through which the psychological mechanisms themselves might operate.

In the following chapter, I turn to Michel Foucault's genealogical method, which I argue better satisfies the explanatory values of a politically realist inquiry. Foucault's genealogies are thoroughly empirical. They also provide a robust account of how our practices are produced, sustained, and transformed. What is more, Foucault's success lies in the fact that his method both draw on historical materials for philosophical reflection and critically examines and reconfigures

[392] For her discussion of complexity of interaction as a feminist value, see Longino, "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy," 47-48.

[393] See R.G. Collingwood, *The Idea of History* (Oxford University Press, 1946), 30.

[394] For his criticism of Williams's genealogy on this topic, see Ian Hacking, "Critical Notice: Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy*" in *Canadian Journal of Philosophy* 34, no. 1, 2004: 142.

the practice of historiography itself. It is to Foucault's deeply empirical genealogy that I now turn.

V.

FOUCAULT'S *AUTOPSY*

In the previous chapter, I argued that Williams's vindicatory genealogy offers greater explanatory power when compared to Queloz's idealized functionalist model because it meets the four criteria for empirical adequacy. Once methods have passed a threshold suitable for empirical adequacy, one might wonder whether further choices between explanations is ultimately a matter of preference.

I argue that we can draw on two further values to adjudicate which empirically adequate explanation we ought to prefer: freedom and complexity of interaction. Explanations that incorporate the value of freedom represent the possibility for the world to be otherwise. Our practices are not fixed to necessary laws of human nature. They are open to revision in ways that are unexpected. The value of freedom therefore provides a necessary critical perspective on our practices. Complexity of interaction prioritizes explanations that treat relationships between entities and processes as co-constitutive rather than unidirectional. Instead of a simplified explanation whereby we account for a given phenomenon by appealing to some one thing that causally determined it, I argue we should prioritize a complex model. This ensures an explanation will speak to the convergence of influences in any given phenomena. Complexity of interaction situates plurality at the level of inquiry itself. Taken together, these two values help us adjudicate the explanations that best satisfy the criteria for empirical adequacy.

In Foucault's genealogies, freedom both animates and unifies the historical and critical dimensions of his empirical accounts. Likewise, Foucault's genealogies foreground the complexity of interaction between the explanatory elements of knowledge and power. Crucially, these contextual values remain responsive to the value of empirical adequacy. Unlike non-concessive utopianism where freedom is not responsive to evidence, Foucault's genealogies reflect what Johanna Oskala describes as "a commitment to a specific historical tradition within which we think about human life and politics."[395] Freedom is not unbounded possibility for Foucault. It is a value for thinking beyond the status quo all while remaining grounded in reality. The historico-critical attitude expressed in genealogical inquiry, Foucault writes, is one "done at

[395] Johanna Oksala, *Foucault on Freedom* (Cambridge University Press, 2009), 190.

the limits of ourselves," one that "puts itself to the test of reality, both to grasp the points where change is possible and desirable, and to determine the precise form this change should take."[396]

We can think of empirical adequacy, freedom, and complexity of interaction as co-regulating values. Empirical adequacy places constraints on what counts as an ambitious yet viable ideal. At the same time, the value of freedom can push against these limits by critically interrogating what is taken to be fixed or stable in human nature and practice. Empirical adequacy also demands explanations capture the complexity of political life, a demand that complexity of interaction helps secure.

To make this argument, I begin with an interpretive and contextual framing of Foucault's historiography. First, I argue that Foucault draws on a broadly Herodotean historiography, which places explanatory priority on events and practices rather than the psychological mental states of historical agents. In addition, Foucault also draws on one of the most fundamental elements of Herodotus's historical methods, namely, the idea of seeing for oneself (*autopsy*). As Foucault draws on *autopsy,* he reworks the term through a critical lens. Seeing (*opsis*) no longer refers simply to what appears before one's eyes, but to the deeper conditions that render seeing and intelligibility possible. Foucault's histories thus concern what lies hidden in plain sight. Likewise, he redefines *autos* as an *ethos* inherited from the Enlightenment. Foucault turns to history in order to confront the authoritative sources of our practices, for the purposes of a possible transgression.

One of the most distinctive features of Herodotus's method is his first-person reflection on that method. This self-reflection makes good sense, given that Herodotus had undertaken an entirely new genre of writing and therefore needed to signal to his audience what they should expect from his historical inquiry. Foucault similarly introduces a new genre of historical and philosophical writing, a fact that helps explain his meta-historical reflections on archaeology as he constructs it. Against this background, I reconstruct Foucault archaeological meta-history in *The Archaeology of Knowledge* before considering its explanatory deficiencies. Drawing on interpretations advanced by Hubert Dreyfus and Paul Rabinow, I argue that archaeology is insufficient insofar as it fails to account for the complexity of interaction between discursive and non-discursive practices. As a result, archaeology's explanatory power falters with respect to

[396] Michel Foucault, "What Is Enlightenment?" in *The Foucault Reader*, ed. Paul Rabinow (Pantheon Books, 1984), 46.

contingency and regularity. I then turn to Foucault's genealogical method to show how its incorporation of strategies and tactics overcomes archaeology's methodological shortcomings. In reading genealogy against archaeology, I argue that genealogical explanation is best suited to politically realist inquiry.

In the final part of my chapter, I offer a direct comparison of Foucault's and Williams's respective genealogies and pay particular attention to contingency and regularity. I argue that we should resist the temptation to explain the suitability of Foucault's genealogy for politically realist inquiry simply in terms of its empirical richness, since the accumulation of empirical detail does not by itself yield a more suitable explanation. What distinguishes competing genealogical explanations, I argue, is how values help us determine which facts count as significant facts and how they in turn comprise a narrative that makes overall sense of them. The value that plays a central role in Foucault's genealogy is complexity of interaction. More specifically, this value enables Foucault to account for regularity in a way that leaves room for contingency. To demonstrate this, I identify two explanatory features in Foucault's genealogy. First, mechanisms, which explain how a process develops, and second, functions, which explain why it does so. I show how the value of complexity of interaction allows Foucault to describe both mechanism and function as constitutive of and immanent to historical practices rather than as determinate or antecedent to them. I then conclude with a brief reflection on the impartiality of Foucault's genealogies.

1. Seeing For Oneself

In *Writing History,* Paul Veyne specifies a methodological problem for the historical disciplines. "The field of history," he writes, is "completely underdetermined, with one exception: everything in it must really have taken place."[397] Historical time does not exist as atoms do but is understood to have been invented by Herodotus as a way of narrating a past we recognize in terms of a determinate past.[398] The historian is tasked with shaping a past with "no natural joints" into a narrative that both makes sense and represents something determinate or real.[399] Williams refers

[397] Paul Veyne, *Writing History: Essay on Epistemology* (Wesleyan University Press, 1984), 15.

[398] Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton University Press, 2002), 163.

[399] Veyne, *Writing History*, 18

to this commitment to historical explanation as a mode of historical truth, which is distinct from the mode of myth or legend.[400]

Foucault borrows a number of methodological tools from historians. Here I focus on one of Foucault's most defining historiographical elements, one traced to Herodotus's *Histories*, written in the fifth century BC as an effort to document the past factually and based on evidence. Herodotus termed his method *historie,* the Greek word for inquiry, and built it on the concept of *autopsy*. In its classical usage, *autopsy* derives from the Greek *autos* ("oneself") and *opsis* ("sight, view"), signifying the act of seeing for oneself. In modern usage, *autopsy* refers to the examination of a body to determine the cause of death, reflecting its broader meaning as a diagnostic tool for understanding the makeup and extent of a disease.[401] This sense of autopsy is also present in some of the earliest historiographies, particularly in Thucydides' *History of the Peloponnesian War.*[402]

Herodotus begins his *Histories* with a meta-narrative (or meta-*historie*) writing, "I, Herodotus of Halicarnassus, am here setting forth my history, that time may not draw the color from what man has brought into being, nor those great and wonderful deeds, manifested by both Greeks and barbarians, fail of their report, and, together with all this, the reason why they fought one another."[403] Herodotus here reflects on the purpose of historical inquiry. It preserves human knowledge, discerns the causes of wars and other major events, and documents *kleos* ("glory"). Undergirding this meta-narrative is history's pragmatic orientation, insofar as it intends to put historical knowledge to use. To know history is to know man, which itself has intrinsic value. This ethos continues to run deep in more contemporary historiography. In *The Idea of History*, British 20th century philosopher, R.G. Collingwood writes, "knowing yourself means knowing what you can do; and since nobody knows what he can do until he tries, the only clue to what

[400] Williams, *Truth and Truthfulness*, 164.

[401] *Oxford English Dictionary*, s.v. "autopsy (*n.*)," July 2023, https://doi.org/10.1093/OED/5307650901.

[402] For a detailed account of the influence of medical language on Thucydides's historiography, and how this informed his ideal of human nature, see Simon Swain, "Man and Medicine in Thucydides," *Arethusa* 27, no. 3 (1994), 303-327.

[403] See Herodotus, *The History*, trans. David Grene (University of Chicago Press, 1987), 1; See also See Nino Luraghi, "Meta-historie: Method and Genre in the Histories," in *The Cambridge Companion to Herodotus*, ed. Carolyn Dewald and John Marincola (Cambirdge University Press, 2006), 76-91.

man can do is what man has done. The value of history, then, is that it teaches us what man has done and thus what man is."[404]

Foucault adopts both senses of *autopsy* in his essay, "Nietzsche, Genealogy, History," and in his subsequent genealogies. Effective history, he tells us, demands a "vast accumulation of material," "relentless erudition" and affirms "a perspectival knowledge [*savoir*]."[405] Effective history records events in those places that we tend to think are without history, for example, in the things "nearest to it—the body, the nervous system, nutrition, digestion, and energies."[406] The historian recognizes the jolts and surprises of history, and diagnoses its illness.[407] In this respect, Foucault writes that "historical sense has more in common with medicine than philosophy."[408] Finally, effective history is not concerned with documenting renown, or *kleos* largely because it does not document men but events and relations. Foucault writes, "no one is responsible for an emergence; no one can glory in it, since it always occurs in the interstice."[409]

Although Foucault incorporates historical methods, he is foremost a philosopher, and philosophy is primarily a reflective activity. It examines its objects of inquiry but considers the conditions that make such an inquiry possible. Accordingly, Foucault's method reflects on the past and the conditions that enable such reflection. Foucault's histories are therefore less concerned with specific details of events—what they were, where, and when they occurred—and more focused on what makes those events possible, and in turn, knowable to the historian. In this sense, Foucault's histories are "impure" as he alters the aim of historical inquiry compared to traditional methods. His histories raise problems and question objects and assumptions we take to be obvious, fixed, or essential.

Seeing in the Foucauldian sense is related to both what appears before one's eyes, such as a symptom, a conflict, a ruin, an archival document and the depth conditions that make seeing and

---

[404] R.G. Collingwood, *The Idea of History* (Oxford University Press, 1946/1956), 10.

[405] Michel Foucault, "Nietzsche, Genealogy, History," in *Essential Works of Foucault Vol. 2*, ed. James D. Faubion (The New Press, 1998), 370, 382.

[406] Foucault, "Nietzsche, Genealogy, History," 381.

[407] Michel Foucault, "Nietzsche, Genealogy, History," 373.

[408] Foucault, "Nietzsche, Genealogy, History," 383.

[409] Foucault, "Nietzsche, Genealogy, History," 377.

intelligibility possible. His histories bring to light what has been hidden, or rather, what lies hidden in plain sight right before our eyes. As Hubert Dreyfus and Paul Rabinow write, Foucault's histories show that "when viewed from the right distance and with the right vision, there is a profound visibility to everything."[410] The critical target of his inquiries is our tendency to habituate ourselves to practices without questioning their origins or purposes. Foucault's approach identifies these problems, traces their emergence with detachment, and makes obvious the dangers the problems lay bare.[411]

In "What Is Enlightenment?" Foucault offers one of his most sustained reflections on the role of *autos* (oneself) within his genealogical method. Genealogy, he writes, is "a type of philosophical interrogation—one that simultaneously problematizes man's relation to the present, man's historical mode of being, and the constitution of the self as an autonomous subject."[412] This critical attitude, he insists, is rooted in the Enlightenment, not merely as a historical moment but as an *ethos*, "a permanent reactivation of an attitude… that could be described as a permanent critique of our historical era."[413] Just as Foucault adds a critical dimension to *opsis* by problematizing what we take to be self-evident and revealing the conditions under which our concepts and practices become intelligible, he likewise subjects *autos* to critical scrutiny. *Autopsy* fixes to itself a critical lens, with freedom one of its central components.

Seeing for oneself is a practice of freedom that aims to diagnose problems for the purposes of their transgression. So conceived, freedom is not a value aimed at emancipation. It is a value defines the kinds of inquiry suited to bringing into view, and critically reflecting on, the authority of our practices. When one sees for oneself, she confronts the authoritative hold of practices, at which point she decides whether she should transgress their limits.

---

[410] Hubert Dreyfus and Paul Rabinow *Michel Foucault: Beyond Structuralism and Hermeneutics* (The University of Chicago Press, 1982), 107.

[411] See Hubert Dreyfus and Paul Rabinow, "What is Maturity?" in *Foucault: A Critical Reader,* ed. David Hoy (Wiley-Blackwell, 1991), 115.

[412] Foucault, "What Is Enlightenment?," 42.

[413] Michel Foucault, "What Is Enlightenment?," 42.

2. Foucault's Meta-History

In Homer's *Iliad*, only the gods can gain knowledge through direct sight: "Tell me now, you Muses who have your homes on Olympos. / For you, who are goddesses, are there, and you know all things, / and we have heard only the rumor of it and know nothing."[414] The gods have knowledge because they see everything, whereas mortals rely on *akoe* (hearing, hearsay). The historian democratizes the ability to see by raising men up to the position of Gods. According to Foucault, history thus "finds its support outside of time and claims to base its judgments on an apocalyptic objectivity."[415]Apocalypse (*apokalupsis)* refers to a revelation, particularly by a god who discloses the truth to some human recipient.[416] The historian thus assumes the role of the gods when writing of events in the past by dissolving the singularity of events into "an ideal continuity—as a theological movement or natural process" and by taking "unusual pains" to erase the elements of his work that reveals his "grounding in a particular time and place."[417]

Thucydides is a clear example of a historian who dissolves the singularity of events into an ideal continuity. His account of the Peloponnesian War seeks to understand an event only because it tells us something about universal patterns of human nature.[418] Collingwood, is highly critical of Thucydides, the hero of Williams's genealogy from the previous chapter. The problem with Thucydides, according to Collingwood, is that, unlike Herodotus who is the proper father of history, Thucydides is the father of psychological history, the chief purpose of which

> is not an event nor yet a complex of events: it is an unchanging rule which governs the relations between events…what chiefly interests Herodotus is the events themselves; what chiefly interests Thucydides is the laws according to which they happen. But these laws are precisely such eternal and unchanging forms as, according to the main trend of Greek thought, are the only knowable things. Thucydides is not the successor of Herodotus in historical thought but the man in whom the historical thought of Herodotus was overlaid and smothered beneath anti-historical motives."[419]

---

[414] Homer, *Iliad*, trans. Richard Lattimore (University of Chicago Press, 1951), 2.484-6.

[415] Foucault, "Nietzsche, Genealogy, History," 379.

[416] *Oxford English Dictionary*, s.v. "apocalypse (*n.*)," December 2024, https://doi.org/10.1093/OED/2355444502.

[417] Foucault, "Nietzsche, Genealogy, History," 382.

[418] Swain, "Man and Medicine in Thucydides," 303.

[419] Collingwood, *The Idea of History* (Oxford University Press, 1946), 30.

In this respect, Foucault's concern with discontinuity and with the singularity of an event, appears, on the face of it at least, to adopt more of a Herodotean approach to history. As with Herodotus, Foucault is concerned with the events themselves. An event is not a decision one makes and consequently no one is responsible for it.[420] Though Foucault does not deny the existence of psychological motivations, he construes them as the effects rather than the source of motivation.[421] In other words, rather than representing truth and power as products of motivations and dispositions, Foucault argues we ought to conceive of these motivations as part of an overall strategic network of "dispositions, maneuvers, tactics, techniques, functionings."[422] Interpretating and explaining events is not a matter of uncovering hidden meanings or motivations, but of recording the history of these system of rules "as they stand for the emergence of different interpretations."[423]

In *The Birth of Biopolitics*, Foucault describes his genealogy as "exactly the opposite of historicism: not, then, questioning universals by using history as a critical method, but starting from the decision that universals do not exist."[424] Foucault's main challenge and contribution to histography is a refined method that disrupts the continuity of man, understood as having both an essentialized body and a fixed set of psychological motivations. Even the body, Foucault maintains, does not obey timeless laws of physiology and as such does not escape the influence of history.[425] As with our psychological motivations, the body is molded by strategies and techniques. Genealogy as "effective" history is hence distinct from "the history of historians in being without constants. Nothing in man—not even his body—is sufficiently stable to serve as the basis for self-recognition or for understanding other men."[426] Foucault thus problematizes traditional history by suspending all presumptions of continuity. When these problematizations are taken together, genealogy differs both in its aim and in its effects. Genealogy, Foucault

[420] Foucault, "Nietzsche, Genealogy, History," 377.

[421] See Dreyfus and Rabinow, *Michel Foucault*, 109.

[422] Foucault, *Discipline and Punish*, 26.

[423] Foucault, "Nietzsche, Genealogy, History," 378-79.

[424] Foucault, *The Birth of Biopolitics: Lectures at the Collège de France, 1978-79* (Palgrave Macmillan, 2008), 3.

[425] Foucault, "Nietzsche, Genealogy, History," 380.

[426] Foucault, "Nietzsche, Genealogy, History," 380.

writes, "will uproot its traditional foundations and relentlessly disrupt its pretended continuity."[427] Reflecting upon one's present disrupts self-understanding to such an extent that even our psychological and corporeal dispositions cannot serve as the basis for self-recognition or continuity.

As a consequence of Foucault's challenges to traditional historiography, genealogy becomes a method that de-sediments rather than expands knowledge. In other words, the reflective, critical thrust of his histories is often more inclined to dismantle knowledge than accumulate or generate it. Of course, there is plenty of knowledge to gain from what Foucault's histories enable us to see. This knowledge does not however lead to greater understanding but introduces "discontinuity into our very being." [428] At the very least, acquiring more knowledge (especially for the sake of knowledge) is not the point of his histories and less so where the value lies. Rather, the point is to make manifest what has been in front of us all along and to challenge what we take to be obviously without history.[429] This might lead to greater knowledge, but it certainly does not have to.

As Gary Gutting notes, much traditional history narrates the inevitability of our present situation, which is given to us by the causes and forces revealed in the narrative.[430] Even though traditional history incorporates contingency into its account, I have demonstrated it does so in a way that ultimately preserves underlying psychological laws that structure events. By contrast, Foucault's historical inquiry focuses on what Johanna Oksala calls the "ontological contingency of the present," which constitutes the bedrock of freedom.[431] Historical inquiry at once strips us of our preconceptions and exposes the unpredictability of the ways of thinking and acting through which the present has emerged. The disruptive nature of Foucault's genealogy thus aims to free us from our inherited assumptions about ways of thinking and being that we take to be fixed and stable.

---

[427] Foucault, "Nietzsche, Genealogy, History," 380.

[428] Foucault, "Nietzsche, Genealogy, History," 380.

[429] Foucault, "Nietzsche, Genealogy, History," 386.

[430] Gary Gutting, "Michel Foucault: A User's Manual," in *Cambridge Companion to Foucault* ed. Gary Gutting (Cambridge University Press, 1994), 10.

[431] Johanna Oksala, *Foucault and Freedom*, 189.

At the same time, genealogy is not concerned with documenting a haphazard chronology of events. Genealogy seeks to make overall sense of why and how those events occurred by rendering them intelligible within a narrative. It may be for this reason that traditional history often relies on the subject as a locus of intelligibility and why it explains historical developments in terms of psychological mechanisms. As I discussed in the previous chapter, Williams's genealogy exemplifies this method insofar as it explains why we adopt the virtues of truth by reference to universal human needs and reactive attitudes, such as shame and honor.

Even in Foucault's historiography there remains a pressure to make overall sense of historical events in a way that is intelligible and coherently structured. As Foucault notes in *Discipline and Punish*, "Is it surprising that prisons resemble factories, schools, barracks, hospitals, which all resemble prisons?"[432] The resemblance of these phenomena can only be made sense of against a historically defined framework of rules and practices, which lends a sense of regularity to contingent events. Perhaps most importantly for the work of freedom, Foucault's genealogies also provide the material through which we critically engage with our present practices and thereby fabricate new ones.

As I mentioned above, one of Herodotus's defining methodological features is his first-person reflection on his own historical practice. As Nino Laraghi notes, there was no real precedent in ancient Greek or Roman historiography for this kind of meta-historical approach.[433] For Laraghi, there are several reasons why Herodotus might have felt it necessary to comment on his method, chief among them, the need to signal to his audience what they should expect from this new genre of historical inquiry.[434] Getting clear on such expectations is especially important when one's method significantly departs from traditional methods. On this point, Laraghi writes, "The need to articulate the rules of a new genre and communicate them to the audience are the most satisfactory explanation for some extremely peculiar statements of meta-*historie*, in which the narrator tells the audience what he is or is not supposed to do—as if inviting them to take part

[432] Foucault, *Discipline and Punish* (Vintage Books, 1975), 228.

[433] See Nino Luraghi, "Meta-historie: Method and Genre in the Histories," in *The Cambridge Companion to Herodotus*, ed. Carolyn Dewald and John Marincola (2006), 76-91.

[434] Laraghi, "Meta-*histore:* Method and genre in the *Histories*," 85.

in a game whose rules they do not yet know exactly, while at the same time showing himself bound by those rules."[435]

I suggest that we read Foucault's own meta-historical reflection in a similar light. Throughout his archaeological and genealogical writings, Foucault intervenes in the first person to reflect on his process of inquiry and to redefine many of the rules of both historiographical and philosophical investigation. By articulating these rules, Foucault constructs a narrative genre that, while suspending continuity and universality with respect to his materials, nonetheless acknowledges the importance of setting expectations for his audience. What follows is a detailed account of Foucault's meta-history, beginning first in archaeology before turning to genealogy. I will then conclude with a discussion of why Foucault's genealogy is best suited to a politically realist inquiry.

3. Rupturing the Past: The Elements of Archaeology

When historical time is paired with something like a discerning attitude, by which we scrutinize the evidence of the remote past as we would of yesterday or last week, we construct a minimal historiography. This method bears Herodotus's mark, whose historical craft did not take the stories of eyewitnesses at face value but subjected them to an entire practice of testimonial discernment, a practice borrowed from the law courts.[436] Herodotus and Thucydides asked the same questions of the witness of events as one would the witness of the courts. Under the auspices of Greek historiography, the memory of the eyewitness serves as our evidence. The historical method aims to elicit their testimony of events into a picture "far fuller and more coherent than any could have volunteered for himself."[437] This has some well-noted drawbacks, most notably, its scope. Had Herodotus reached too far into the past and recounted something akin to historical hearsay, the historical method would have retreated into the mode of myth or legend and would have ceased to capture something significant about the past as *our* past. An effective historical method would require, for the time being, a myopic perspective.

[435] Laraghi, "Meta-*histore:* Method and genre in the *Histories*," 85.

[436] See Collingwood, *The Idea of History*, 17-20.

[437] Collingwood, *The Idea of History*, 25.

The historical practices ushered in by Herodotus and Thucydides have yet to wane. The historian, according to Marc Bloch, "does not think of the human in the abstract. His thoughts breathe freely in the air of the climate of time."[438] Historians seek definitive answers by asking definitive questions through programmatic research and the careful scrutiny of evidence.[439] They also grapple with similar methodological problems, the first of which is history's ability to sufficiently capture the concrete reality of the past, particularly when the evidence of the past is always indirect.[440] History's claim to empiricity is well-founded but nevertheless insecure, since the historian always observes the past through the eyes of others, and so finds herself in a similar predicament as a lawyer when cross-examining a witness. Despite this methodological fragility, historians still assert we can apprehend a sense of the past with confidence.

One way historians lend greater stability to their narratives is by constructing them from archival documents. The historian's craft today is distinctly marked by the archive. Bloch sets up some preliminary methodological problems archival research might consider. For one, the idea that forgeries do not inherently lie or deceive reveals that we are not always clear on the difference between what is real or false. Under certain conditions, fraud is counted as an authentic piece of evidence through which we, Bloch describes, seek out "the imposter behind the imposture."[441] In exceptional cases, documents may have been forged with the intention to replicate real papers that have been lost.[442] Yet, if we admit forgeries of this sort into the historical narrative, then we should also consider the more fringe cases, such as when the historian conjures the remains of the past precisely from documents that never existed, that never made their way into the archive, or that were destroyed and never replaced. Going this far, however, might take us out of the archive altogether and back into the mode of myth.

---

[438] Marc Bloch, *The Historian's Craft: Reflections on the Nature and Uses of History and the Techniques and Methods of Those Who Write It* (Vintage Books, 1949), 27.

[439] Collingwood, *The Idea of History*, 18.

[440] Bloch, *The Historian's Craft*, 48.

[441]Bloch, *The Historian's Craft*, 93.

[442]Bloch, *The Historian's Craft*, 91.

Another related problem Bloch considers is about the construction of the archive itself. How do the various archival guides, such as library catalogues, museum indexes, and bibliographies make their way into the archive in the first place? Bloch writes,

> despite what the beginners sometimes seem to imagine, documents do not suddenly materialize, in one place or another, as if by some mysterious decree of the gods. Their presence or absence in the depths of this archive or that library are due to human causes which by no means elude analysis...[E]very historical book worthy of the name ought to include a chapter, or if one prefers, a series of paragraphs inserted at turning points in the development, which might be entitled: 'How can I know what I am about to say?[443]

Bloch recognizes that the archive is shaped by history, and as such, is constitutive of contingency both in its emergence as a repository for nominal inquiry and in the various and obscure way things make themselves into that repository. What he does not provide, and perhaps cannot owing to his own methodological limitations, is a fuller account of that history. In other words, Bloch is unequipped with the methodological tools that would allow him to explain how historical knowledge, and knowledge in general, is constituted in the first place.

Rather than ask, "How can I know what I am about to say?" we might instead ask, "What allows the archive to appear before me? What are the conditions that make the materials of the archive intelligible?" Answering these questions requires a critical perspective, one that explains the conditions of knowledge itself. For a critical approach to remain historical, however, the explanation must not rely on immutable laws, essential essences, or similar foundational concepts. In other words, it must be both sufficiently critical *and* historical.

In his preface to *The Birth of the Clinic*, Foucault describes his inquiry as "deliberately both historical and critical, in that it is concerned—outside all prescriptive intent—with determining the conditions of possibility of medical experience in modern times."[444] Foucault seeks to uncover the historical rules that have governed a particular mode of experience and to do so in a way that walks the line between, as Béatrice Han puts it, a precritical realism and an idealistic transcendentalism.[445] In other words, he seeks to give an account of these rules that is neither a

[443]Bloch, *The Historian's Craft,* 71.

[444] Michel Foucault, *Birth of the Clinic* (Vintage Press, 1963), xix.

[445] Béatrice Han, *Foucault's Critical Project: Between the Transcendental and the Historical* (Stanford University Press, 2002), 41.

mere chronology of empirical events nor an atemporal structure that governs the space of transformations between *epistemes*. As he will later describe in *The Archaeology of Knowledge*, Foucault seeks to account for specific regularities.[446]

Foucault's definitive text on the archaeological method is at once a final rapprochement of his work in *The History of Madness*, *The Birth of the Clinic*, and *The Order of Things*, and an attempt to establish his own method as distinct from traditional history. Foucault's commentary on archaeology first removes spurious groups and categories of regularity to which historians and philosophers often assign special methodological status: periods, traditions, influence, development and evolution, forms of genre, the book, or the *oeuvre*.[447] When carrying out historical inquiry in the archive, these groupings seem like natural places to begin. They prove less stable, however, once they are subjected to critical scrutiny. Take, for instance, the book: is the material unity of a book the same as an anthology of poems, or a collection of posthumous fragments? Even more difficult is the *oeuvre*: we assume it is a collection of texts we designate under the header of a proper name, but what do we make of works published under a pseudonym, or the unfinished drafts and collections of notes in a notebook? Foucault writes, "As soon as one questions that unity it loses its self-evidence; it indicates itself, constructs itself, only on the basis of a complex field of discourse."[448] Foucault's archaeology throws these categories into doubt and in so doing deprives us of continuity, dissipates temporal identity, to breaks "the thread of transcendental teleologies."[449]

It is worth clarifying here that Foucault's dismantling of traditional historiography does not amount to a full-scale rejection of categories and kinds as such. As we will see, Foucault disassembles these ostensible unities only to reassemble with a critical perspective.[450] He asserts

---

[446] See Michel Foucault, "Discursive Formations," in *The Archaeology of Knowledge* (Vintage Books, 1969), 31-39.

[447] These are all mentioned in chapter "The Unities of Discourse" in *Archaeology of Knowledge,* 21-30.

[448] Michel Foucault, *The Archaeology of Knowledge* (Vintage Books, 1969), 23.

[449] Foucault, *The Archaeology of Knowledge*, 131.

[450] In *Foucault's Archaeology of Scientific Reason*, Gutting writes: "This insistence on discontinuity has led many of Foucault's readers to think that his archaeological approach allows no place for gradual transformation or continuous developments, that it sees the history of thought as a series of quantum leaps from one self-contained discursive formation to another. (Compare the similar reaction of readers of Thomas Kuhn to his notion of incommensurability of paradigms.)" (247,); also see Ian Hacking "Michel Foucault's Immature Science," in *Historical Ontology* (Harvard University Press, 2002), 87-98.

that a method that unifies elements "must not be rejected definitively of course, but the tranquility with which they are accepted must be disturbed…. what we must do, in fact, is to tear away from them their virtual self-evidence, and to free the problems that they pose."[451] In the penultimate chapter to *The Archaeology of Knowledge*, Foucault reminds the reader that in suspending the "theme that succession is an absolute," she must also take care not to freeze time once and for all, less the archaeologist re-instantiate a new kind of atemporal structure, where "the possibility of a historical description disappears."[452] Archaeology should not invert "the usual values" of continuity and discontinuity," but show how the continuous is formed "in accordance with the same conditions and the same rules as dispersion" and how it comes about through a specific discursive practice.[453] In other words, archaeology strengthens our understanding of how discursive practices persist through time against the backdrop of discontinuity. The historian's craft thus becomes one of fashioning the innumerable practices of the past into intelligible sequences, which Foucault often refers to as *positivities of discourse*; for such intelligibility enables historiography to build a condition of reality for these practices.[454]

After suspending certainty in the unities most historians unreflectively take as given, Foucault reconstructs a historiography of dispersed elements that are nevertheless identified through specific regularities. As is typical in historical inquiry, Foucault begins with the documents of the archive, which are, in effect, collections of what he calls *statements*. From there, historical inquiry often proceeds by interpreting and reconstructing what lies beneath these documents, whether by uncovering the thought and inner life of the author or by treating concepts as objective linguistic data, whose continuity is then revealed through analysis. This, however, already takes too much for granted. Foucault instead treats the statement on its own terms.

The statement as a starting point of analysis would not have been altogether new for philosophy. Grammar, for instance defines the conditions that render statements meaningful. Logic specifies what can be added to statements without contradiction. Both analyze the

[451] Foucault, *The Archaeology of Knowledge*, 25.

[452] Foucault, *The Archaeology of Knowledge,* 166.

[453] Foucault, *The Archaeology of Knowledge,* 175.

[454] Foucault, *The Archaeology of Knowledge*, 125, 127.

statement in its own right. The mistake of grammar and logic, according to Foucault, is that they treat statements as isolable objects as opposed to as enunciative functions, whose meaning and intelligibility changes with history. Treating the statement historically involves recognizing "there is no statement in general, no free, neutral, independent statement; but a statement always belongs to a series or a whole, always plays a role among other statements, deriving support from them and distinguishing itself from them; it is always part of a network of statements, in which it has a role, however minimal it may be, to play."[455] Foucault claims this network of statements is defined as a referential (i.e., a field of emergence), a subject (i.e., someone or something that emits the statement), associated field (i.e., the other statements with which coexists), and materiality (i.e., a substance, a support, a place, a date).[456] A statement's status as intelligible and meaningful is thus defined by how it functions with respect to all of these domains, which is also what gives the statement its historical dimension.

Once Foucault has found a starting point for historical inquiry that is neither static nor fully dispersed, he begins to group statements into *discursive formations*, which are comprised of four elements: its object (e.g., madness, delinquency, sexuality), enunciative modality (e.g., the authoritative status of that who utters the statement), its concept (e.g., its recurrent elements in various types of discourse), and its themes or strategies (e.g., its theoretical viewpoint, such as biology, physics, grammar, etc.,).[457] Crucially, we cannot describe a discursive formation in terms of any one element. Discursive formations take shape within and through "a system of dispersion" that divides the elements into seemingly "incompatible themes" that are nevertheless marked by some regularity.[458] This system forms the background conditions that unify and govern statements in their discursive formations through an anonymous system of rules. In making a statement, a speaker therefore takes up a position that has already been defined by the rules of a given discursive formation "without reference to a cogito."[459] Foucault concludes his description of his archaeological elements with four propositions, which I paraphrase here:

---

[455] Foucault, *The Archaeology of Knowledge*, 99.

[456] See Foucault, "The Enunciative Function" in *The Archaeology of Knowledge*, 88-105.

[457] For this schematization, see Foucault, "Discursive Formations," in *The Archaeology of Knowledge,* 31-39.

[458] Foucault, *The Archaeology of Knowledge,* 37.

[459] Foucault, *The Archaeology of Knowledge*, 122.

1. There is a correlative relationship between statements and discursive formations: a discursive formation consists of statements, and a statement acquires meaning, truth, and intelligibility only within a discursive formation.
2. A statement's position within a discursive formation is inseparable from its adherence to the rules of a discursive formation. In other words, discursive formations are *expressions* of the rules that constitute them.
3. Discourse, understood as a set of statements within a discursive formation is neither a formal nor ideal unity, but is thoroughly historical.
4. A discursive practice consists of a body of anonymous, historical rules that determine the conditions of operation of the statement. These discursive practices give reality to statement *as expressions* of the rule-governed system.[460]

Even when archaeology describes discursive practices as expressions of the rules of formation, the method must still explanatorily connect the rules to their corresponding practices. One way to understand why Foucault requires this additional connection is that he seeks to explain rules from an internal point of view as opposed to an external one. In normative theory, this distinction accounts for the motivations individuals have toward rules and the intelligibility of discursive practices. Whereas an external point of view identifies patterns of behavior and presents rules as mere descriptions of that behavior, an internal account assumes the stronger position that participants accept and use rules as normative guides for action. Rules, from an internal point of view provide individuals with reasons for action and so are said to have normative force. I draw this distinction from H.L.A Hart's *Concept of Law*, which describes the normative force of legal rules in terms of a participant's acceptance of those rules as standards guiding conduct rather than backed by sanctions or described as observable regularities in behavior. What we acquire from Hart's account is that an internal point of view need not rest of an account of psychological motivation, but simply a shared practice of rule-following.[461] Like Hart, Foucault explains discursive practices "in accordance with specific regularities" that guide practices.[462] He must now formulate a principle of determination for action that, while normative, is both anonymous and "internal" in Hart's sense. Foucault calls this principle the historical *a priori*.

---

[460] See Foucault, *The Archaeology of Knowledge*, 117.

[461] My description here is less connected to Williams's reason internalism, which states that genuine reasons for action must connect to the agent's existing motivational states rather than through purely external pressures. See H.L.A Hart, *The Concept of Law* (Oxford University Press, 1961); Bernard Williams, "Internal and External Reasons," in *Moral Luck* (Cambridge University Press, 1981), 101-113.

[462] Foucault, *The Archaeology of Knowledge*, 128.

The term "historical *a priori*" is nebulous. It is implicitly present in *The Birth of the Clinic*, explicitly in *The Order of Things*, and most extensively discussed in *The Archaeology of Knowledge*.[463] Nowhere does Foucault lend a stable definition to the term. In general, the historical *a priori* functions as the explanatory mechanism that unifies statements and so enables archaeologists to assert that subjects are "talking about 'the same thing'" insofar as their practices are motivated by the same reasons.[464] What remains clear throughout Foucault's work is his invocation of Kant's formal *a priori*, which establishes the conditions of validity for judgments. The historical *a priori*, by contrast, serves as "the condition of reality for statements" that are actually said.[465] This *a priori* is "a purely empirical figure. It is not self-evident or necessary in the way we typically think of *a priori* conditions, but neither is it "an absolutely extrinsic contingence."[466] It is historically pliable insofar as the rules that characterize a discursive practice are themselves "caught up in the things they connect."[467] We might say that the historical *a priori* defines the limits of intelligibly for the rules that govern discursive formations, and by extension, for the reasons that mark the standards of conduct for a given period.

Let us take stock. Foucault's archaeology has done away with the often-presumed unities of historiography, such as books, oeuvres, and traditions, and replaces them with new historiographical elements: groups of statements, discursive formations, discursive practices, and rules of regularity. He then introduces the historical *a priori* as a mechanism that unifies these formations, and it is through conformity to a given *a priori* that a discursive practice can be recognized as meaningful, truthful, and intelligible within that system. Moreover, the standard of conduct established by this *a priori* is adopted as a reason for action by those operating within its bounds.

---

[463] For example, in *The Birth of the Clinic*, Foucault writes, "Medicine made its appearance as a clinical science in conditions which define, together with its historical possibility, the domain of its experience and the structure of its rationality. They form its concrete *a priori,* which is now possible to uncover, perhaps because a new experience of disease is coming into being that will make possible a historical and critical understanding of the old experience" (xv).

[464] Foucault, *The Archaeology of Knowledge*, 126.

[465] Foucault, *The Archaeology of Knowledge*, 127.

[466] Foucault, *The Archaeology of Knowledge*, 128.

[467] Foucault, *The Archaeology of Knowledge*, 127.

With this framework in place, Foucault concludes his historiographical reconstruction with the archive, the traditional starting point of historical inquiry. However, archive of archaeology markedly differs from the brick-and-mortar institutions implied in Bloch's *Historian's Craft*. For Foucault, the archive stands as an expression of the historical *a priori* and is recast as a system of discursive practices. It is "the law of what can be said" and apprehended on the "particular level" at which such practices persist, transform, or fade away.[468] In more concrete practices, the archaeologist must draw on this archive in order to reflect on and make sense of the historical *a priori* and its constitutive rules of formation. In this regard, archaeology is an inquiry that can be undertaken only through an active engagement with the empirical elements of the archive.

Foucault's suspension of conventional historiographical elements and his reconstitution of the archive as the site where the historical *a priori* gets expressed provides a more expansive methodology and enables an answer to Bloch's question above. No longer limited to reflecting on the statements contained within the document, historical inquiry can now distinguish between the intelligible statements of the archive and the epistemological conditions that rendered those statements intelligible in the first place. This distinction is best understood as a difference between *connaissance*, which refers to the particular formalized bodies of knowledge (such as history, biology, or physics), and *savoir*, which refers to the underlying depth conditions that structure and unify the formalized bodies of knowledge. In other words, archaeology opens up knowledge at the level of *savoir*, which specifies how statements are subject to normative constraints and how that regulation produces distinctive patterns of intelligibility for the domains that make up *connaissance*. However, although every level of *connaissance* contains its corresponding *savoir*—"there is no knowledge (*savoir*) without a particular discursive practice"—a specific *savoir* does not necessarily give rise to *connaissance*.[469] *Savoir* instead forms "the precondition of what is later revealed and which later functions as an item of knowledge (*connaissance*) or an illusion, an accepted truth or an exposed error, a definitive acquisition or an obstacle.*"*[470] *Savoir* is not an epistemological given (*donnée*), a "pre-knowledge," or an "archaic stage" of immature knowledge, It is set of discursive practices that

[468] Foucault, *The Archaeology of Knowledge,* 129-130.

[469] Foucault, *The Archaeology of Knowledge*, 183.

[470] Foucault, *The Archaeology of Knowledge*, 182.

define the domain of intelligibility and normative constraints of a specific *connaissance*.[471] In other words, *savoir*, comprised of discursive formations and practices, defines at its limits what counts as a serious statement, and what would amount to a domain being either true or false.

With the archive now serving as the stable starting point, the archaeologist can apprehend discursive practices at the level of *savoir*. In other words, archaeology's aim is to specify *savoir* so as to foreground the background conditions upon which the rules of discourse operate without an identifiable subject. The reader of an archaeological explanation can, in turn, recognize and critically assess the discursive practices that emerge as historically constituted forms of rationality. This articulation of archaeology as system of rules at the level of *savoir* underscores its critical thrust. It also fuels skepticism among those who question its historical merits.

4. Is Savoir Sufficient?

Foucault's archaeology has faced sustained critiques for granting autonomy to the rules of formation and the corresponding normative force they exert on our practices. Some characterize archaeology as a methodological failure.[472] Others take a more reconciliatory approach by describing archaeology as facing certain limitations, in the way that most methods do.[473] Fewer attempt to vindicate archaeology.[474]

Archaeology aims to uncover the rules of *savoir* that regulate domains of knowledge (*connaissance*) independently of the consciousness of those who follow them. With this epistemic level now apprehended, archaeology can account for normative patterns without making recourse to subjective determination. It also presents a paradox regarding the normative force of rules, which appear to operate both non-intentionally and from an internal point of view. For Hubert Dreyfus and Paul Rabinow, it is unclear how rules are both efficacious, meaning the participants accept and use rules as normative guides for their practices, and yet unknown to

[471] See Gary Gutting, *Michel Foucault's Archaeology of Scientific Reason*, 251.

[472] Dreyfus and Rabinow, *Michel Foucault*, Chapter Four; see also Béatrice Han, *Foucault's Critical Project,* 38-74; Gutting, *Michel Foucault's Archaeology of Scientific Reason*, 261-288.

[473] See Colin Koopman, "Critical Historiography: Politics, Philosophy & Problematization" in *Genealogy as Critique* and Arnold Davidson, "On Epistemology and Archeology: From Canguilhem to Foucault," in *The Emergence of Sexuality: Historical Epistemology and the Formation of Concepts* (Harvard University Press, 2001).

[474] Tuomo Tiisala, *Power and Freedom in the Space of Reasons: Elaborating Foucault's Pragmatism* (Routledge Press, 2024).

those very participants. According to Dreyfus and Rabinow, Foucault must choose between two views: an external point of view, in which rules function as descriptive regularities that people act in accordance with (as a kind of behaviorism), or an internal point of view, in which rules operate prescriptively as normative forces that people consciously follow (i.e., the rules function as reasons for action).[475] On an explanatory level, it seems incoherent to say that both occur. After all, how can people follow rules from an internal viewpoint all while unaware of their existence? Foucault must therefore choose between an internal viewpoint, thereby reinscribing psychological motivations back into his explanation, or an external viewpoint, thereby forfeiting the ability to argue that statements obey the rules of regularity in a causally efficacious way.

Foucault's ambivalence on this point leads Dreyfus and Rabinow to declare archaeology a failure insofar as it lacks a thoroughly historical account of the rules of *savoir*. If Foucault aims to offer a critical account without presupposing a subjective determination of rules, then he would be best served by a structuralist approach that locates the causal efficacy of rules in mechanisms akin to physical laws operating in the brain, as seen in the structuralism of Noam Chomsky and Claude Lévis-Strauss.[476] However, structural explanations of this ilk lose their historical sense twice over. By grounding the laws in some kind of physical manifestation, Foucault fails to represent the rules as historically variable, that is, contingent and thereby subject to change rather than self-evident (*a priori)*. Second, even if Foucault were to reject such a reductionism, the rules would still lack historical explanatory power. Archaeology can tell us *that* rules conditioned past discursive formations but not how they did so or how one formation transformed into another. To claim that the rules of formation had normative force, Foucault would have to concede that social practices are simply the outcome of linguistic rules. It is not difficult to see how reductionist this would be. As David Hoy observes, archaeology bears "vestiges of idealism."[477] Although the rules of *savoir* allow for a critical explanation, they do not, according to Dreyfus and Rabinow, allow for a historical explanation.

One can sense Foucault's desire to avoid the trappings of linguistic idealism during his archaeological period. This is clearest in his ambivalent handling of the relationship between

[475] Dreyfus and Rabinow, *Michel Foucault*, 81.

[476] Dreyfus and Rabinow, *Michel Foucault*, 82-83.

[477] David Hoy, "Foucault's Slalom," in *The London Review of Books* 4, no. 20 (November 4, 1982).

discursive and non-discursive practices. In *The Archaeology of Knowledge*, Foucault prioritizes discursive formations but also acknowledges the role non-discursive practices have in shaping the rules of *savoir*. He characterizes archaeology as that which "individualizes and describes discursive formations," and notes that it "reveals relations between discursive formations and non-discursive domains (institutions, political events, economic practices, and processes)."[478] His most explicit discussion of the relationship appears in his chapter "The Formation of Strategies." There he frames his archaeological projects in terms of the elements they track and their respective rules of formation: *The History of Madness* focused on objects, *The Birth of the Clinic* on enunciative modalities, and *The Order of Things* on concepts.

Strategies account for why certain formations between objects, enunciative modalities, and concepts emerged while others did not. Formally, they explain how particular formations of *savoir* take shape without resorting to transcendental explanations. Formations arise, according to Foucault, through what he calls *authorities,* one type of which are the rules of regularity; however, they also emerge through "the *function* that the discourse under study must carry out *in a field of non-discursive practices*."[479] This suggests formations take shape within a "system of different strategies that are deployed in it," which encompasses discursive practices, non-discursive practices, as well as processes of appropriation and positions of desire.[480] Only by situating the rules of *savoir* within this broader strategic alignment can we fully understand them in historical terms.

Though strategies appear paramount to understanding the rules of *savoir*, Foucault recognizes that his analysis has "scarcely touched on their formation," and leaves this topic for future studies.[481] Notwithstanding his acknowledgement of a fuller analysis of strategies and of the interaction between discursive and nondiscursive practices, Foucault highlights the autonomy of discursive formations within archaeology: "A statement belongs to a discursive formation as a sentence belongs to a text, and a proposition to a deductive whole. But whereas the regularity of a sentence is defined by the laws of a language (*langue*), and that of a proposition by the laws of

[478] Foucault, *The Archaeology of Knowledge*, 157, 162.

[479] Foucault writes that this authority is also characterized by "the rules and processes of appropriation of discourse" as well as the "*possible positions of desire in relation to discourse*" (*AK*, 68).

[480] Foucault, *The Archaeology of Knowledge,* 68.

[481] Foucault, *The Archaeology of Knowledge*, 65.

a logic, *the regularity of statements is defined by the discursive formation itself*. The fact of its belonging to a discursive formation and the laws that govern it are one and the same thing."[482] The relationship between a discursive formation and the rules that regulate it is "established correlatively," meaning that the regularity of a statement, which conditions discursive formations, is itself conditioned by that very formation.[483] This relationship thereby remains fully autonomous from the non-discursive, resulting in what Dreyfus and Rabinow describe as "a strange notion of regularities which regulate themselves."[484]

Instead of classifying archaeology as a methodological failure, some scholars opt for a reconciliatory approach.[485] Arnold Davidson, for example, argues we can isolate the normative rules of discursive practices without taking the stronger position that one can *only* analyze the relations between discursive practices. Put differently, archaeology elucidates interdiscursive relations of *savoir* in their own specificity but does not preclude a larger investigation into the extradiscursive relations between the discursive and nondiscursive.[486] From genealogy, we gain the analytic of power, which incorporates historical practices that cannot be reduced to mere assertations. It also infuses historical inquiry with a dynamic dimension, offering a framework for making sense of both causation and extradiscursive elements.[487]

Alternatively, one might choose to altogether vindicate archaeology, as seen in recent work by Tuomo Tiisala.[488] His defense of archaeology is directly positioned against Dreyfus and Rabinow's critique, who according to Tiisala, misinterpret the core feature of archaeology as a pragmatist elaboration of Kantian thought. Understood in this way, archaeology offers a coherent reconciliation of how Foucault can have a nonsubjective yet internal view of rules of *savoir*, or account of how these rules can be both implicit and efficacious.

---

[482] Foucault, *The Archaeology of Knowledge,* 116.

[483] Foucault, *The Archaeology of Knowledge,* 116.

[484] Dreyfus and Rabinow, *Michel Foucault*, 84.

[485] Davison, *The Emergence of Sexuality*, 205.

[486] Davidson, *The Emergence of Sexuality*, 205.

[487] See also Colin Koopman, who builds on Davidson's reading in "Foucault's Historiographical Expansion: Adding Genealogy to Archaeology," *Journal of the Philosophy of History* 2 no. 3 (2008), 338-362.

[488] See Tiisala, *Power and Freedom in the Space of Reasons*.

The idea that knowledge can be grasped in practical rather than representational terms is central to Tiisala's pragmatic reading of the rules of *savoir*. Such knowledge is both non-intentional and non-accidental, as individuals conform to rules without explicitly articulating them as a set of propositions. Instead, normative rules are taught and followed *non-linguistically* as a practical ability. According to this view, one can coherently conform to a rule from an internal point of view without representing it as such.[489] Foucault's distinction between *connaissance* and *savoir* is thus understood as a distinction between theoretical and practical knowledge. Both guide our discursive practices, but while *connaissance* consists of clearly articulated rules, *savoir* functions as a system of dispositional rules, exerting normative constraints "behind our backs," as Tiisala puts it.[490]

Tiisala identifies a second misunderstanding in Dreyfus and Rabinow's interpretation of the rules of *savoir* and the corresponding discursive and nondiscursive practices. The received reading takes this distinction in a broad sense, according to which discursive practices have a conceptual structure, while nondiscursive practices do not. However, Tiisala argues that Foucault intends a narrower distinction under which discursive and nondiscursive practices are both conceptually constituted, in part because Foucault holds that human experience is necessarily conceptual. To support this claim, Tiisala cites a 1981 interview in which Foucault states: "Thought, it exists, well beyond the below systems and edifies of discourse. It is something that is often hidden but always animates everyday behavior. There is always a little thought even in silent habits."[491] According to this narrower distinction, discursive practices are conceptually determined and produce statements through representational forms, while nondiscursive practices are also conceptually determined and produce statements through socially coordinated dispositions. *Savoir* functions as the "dispositional bedrock" for a system of thought and structures the use of concepts in both theoretical and practical reasoning. Because Foucault is concerned only with conceptually determined human experience, and because *savoir* encompasses the full range of conceptual discursive practices, Tiisala concludes that the rules of *savoir* are coherently autonomous: they are "created, sustained, and sometimes transformed

[489] Tiisala, *Power and Freedom in the Space of Reasons*, 18.

[490] Tiisala, *Power and Freedom in the Space of Reasons*, 25.

[491] Michel Foucault, "Est-il donc important de pesnser?" in *Dits et écrits II*, 1976-1988, ed. Daniel Defert and Fraçois Ewald (Paris: Gallimard, 2001), 999; Quoted in Tiisala, *Power and Freedom in the Space of Reasons*, 68.

through the very activity of making statements without representing the rules as such."[492] *Savoir* thus encompasses an entire system of thought, one that, according to Tiisala's reading, constitutes *the* form of action.[493]

If all historical practices are ultimately conditioned by *savoir*, then Davidson's articulation of genealogy as a method for tracking interdiscursive relations of *savoir* can no longer hold. On Tiisala's account, this is not because archaeology is simply reinscribed into genealogy, but rather because archaeological analysis forms the foundation from which genealogical inquiry emerges. For Tiisala, *savoir* "is an object common between the two lines of analysis" of archaeology and genealogy insofar as they focus "on two complementary tasks of disclosing different configurations of *savoir* and accounting for their historical transformations."[494] Both the production of truth claims, as conditions for knowledge, and the field of possible action, in which power operates, are situated at the implicit level of *savoir*.

It may be that Dreyfus and Rabinow find archaeology insufficient because they value explanations that foreground the complexity of interaction. By complex interaction, I mean that the processes that explain how and why individuals adopt certain practices involve multiple elements rather than a single explanatory features. In addition, these elements are described as reciprocal and co-constitutive rather than one causally determining the other. Complexity of interaction therefore presupposes a form of methodological pluralism, understood here as a method that reflects distinct types of casually efficacious entities rather than just one. As I have argued, idealized theories are abstractions that reduce the world's complexity by isolating certain features at the expense of others. Although such idealization may be appropriate for some explanatory purposes it is ill-suited to explaining how or why individuals adopt certain rules and practices.

Despite Tiisala's claim that his interpretation foregrounds contingency and the historical variability of discursive practices, his account offers little by way of explaining how such practices come to be taken up in a manner that does not simply reduce them to a form of linguistic idealism. It is one thing to argue that a method can analytically distinguish between

[492] Tiisala, *Power and Freedom in the Space of Reasons*, 59.

[493] Tiisala, *Power and Freedom in the Space of Reasons*, 87.

[494] Tiisala, *Power and Freedom in the Space of Reasons*, 45.

normative rules of discursive and nondiscursive practices. It is another thing to claim that social practices and domains of knowledge are merely products of linguistic rules. General reflection on our experience suggests that there is a wide range of practices that lack an underlying discursive structure. Characterizing experience as necessarily conceptual, and human conduct as conceptually determined, not only diminishes the plane of contingency by overdetermining our reasons for action but also fails to account for the complexity of the processes through which individuals come to adopt particular modes of conduct.

Tiisala's interpretation ascribes to Foucault many of the ahistorical tendencies characteristic of a Thucydidean outlook. Like Thucydides, Tiisala presents human conduct as historically variable while ultimately grounding that variability in our nature as human beings, whose capacity for action is "socially aquired through linguistic training."[495] Even practices we feel are without conceptual form, such as gestures and routines, ultimately admit of a conceptual nature. Tiisala's misunderstanding of Foucault's Herodotean spirit reveals the danger of single-axis historical explanations. Conversely it demonstrates why explanations that privilege mutual interaction yield more realistic accounts of human conduct.

It is therefore difficult to reconcile Tiisala's account with Foucault's more Herodotean approach to history, which I described above as focusing on contingent events rather than on underlying psychological structure that defines human action. Events, according to Foucault, are much more accurately described as a relation of war and struggle than a system or formation of thought. In one interview Foucault explicitly contrasts the model of war against a model of language, writing:

> The problem is at once to distinguish among events, to differentiate the networks and levels to which they belong, and to reconstitute the lines along which they are connected and engender one another. From this follows a refusal of analyses couched in terms of the symbolic field or the domain of signifying structures, and a recourse to analyses in terms of the genealogy of relations of force, strategic developments, and tactics. Here I believe one's point of reference should not be to the great model of language (*langue*) and signs, but to that of war and battle. The history which bears and determines us has the form of a war rather than that of a language: relations of power, not relations of meaning. History has no 'meaning', though this is not to say that it is absurd or incoherent. On the contrary, it is intelligible and should be susceptible of analysis down to the smallest detail—but this is in accordance with the intelligibility of struggles, of strategies

[495] Tiisala, *Power and Freedom in the Space of Reasons*, 109.

> and tactics. Neither the dialectic, as logic of contradictions, nor semiotics, as the structure of communication, can account for the intrinsic intelligibility of conflicts.[496]

We see here that the totality of power relations does not make up a meaningful system. Even if meaningful actions occur within these alignments, we can only interpret them against a particular configuration of forces. Indeed, the entire idea of war is that it is senseless. Although rules may be operative within war and we do our best to make historical sense of its events, war itself is often a matter of tactics and strategies.[497] To consider how we might make sense of strategies and tactics, without reducing them to a conceptual structure, we must turn to genealogy.

5. Emergence and Descent: The Elements of Genealogy

Now that I have made the limitations of the archaeological method clear, I turn to argue that the genealogical method is best understood as corrective or additive rather than as an outright rejection of the archaeology. More specifically, Foucault's genealogy amends two explanatory deficiencies of his earlier works.

First, although archaeology is diagnostic, it offers limited practical value for apprehending our own archive and discursive formations. Although it accounts for temporal discontinuities between present and past systems of intelligibility, it does not provide an account of how our current discursive systems emerged or continue to develop from those past formations. Foucault's genealogy introduces methodological tools for a practical critique of our own discursive formations and in a way that is responsive to present practical problems. Take, for instance, how Foucault frames the central problem of *Discipline and Punish*. There, he considers how punishment and the prison itself is sustained by political technologies of the body. At the time of his writing, prisoners in France were revolting against the conditions of confinement, against medical treatment, and against educational services, among others. He points to the increase in prison revolts around the world and the paradoxical nature of "their aims, their slogans, and the way they took place."[498] Setting aside the content of Foucault's genealogy of

---

[496] Foucault, "Truth and Power," in *Power/Knowledge: Selected Interviews & Other Writings 1972-1977* (Vintage Press: 1977/1980), 114.

[497] See Joseph Rouse, "Power/Knowledge," in *The Cambridge Companion to Foucault*, ed. Gary Gutting (Cambridge University Press, 2006), 95-122.

[498] Foucault, *Discipline and Punish*, 31.

punishment for now, the key point here is that he focuses on the "minute material details" sustaining contemporary discursive practices.[499] As with archaeology, Foucault still refutes the historiographical tendency to posit continuities such as the subject or body. The difference, however, is that genealogical inquiry traces practices in terms of the problematizations that make them possible.[500] Genealogy is diagnostic in a stronger, practical sense since it identifies how practices have developed over time and opens up strategic avenues for acting, thinking, and being in the present.

A second amendment genealogy makes to archaeology concerns its explanatory force. As I noted above, archaeology de-sediments the rules that condition past discursive systems, but says very little, if anything, about the normative efficacy of these systems or how they transformed into new ones. Consider by contrast how Foucault begins *Discipline and Punish* by juxtaposing the public execution against the timetable, each of which exemplify distinct penal techniques that punish different crimes and produce different types of subjects.[501] In this case, genealogy reveals both the distinct discursive systems at work and the transformation of the public execution into the timetable less than a century later. Part of how Foucault makes sense of the transformation of discursive formations is by analyzing forms of knowledge in tandem with forms of power.[502] From a genealogical point of view, a statement becomes intelligible only within a network of strategies. So conceived, practices enact the discourse and constitute the thread that binds knowledge to its technical reality.[503]

To grasp this positive account of power, however, it is first important to understand what it is not. Power is not a set of institutions that enforce obedience through a top-down structure, nor is it a juridical system of rule-based subjugation or domination of one group by another.[504] Although power may manifest through such relations, it does so only in historically specific configurations. It cannot be seized, acquired, or located external to other relations. It is neither a

[499] Foucault, *Discipline and Punish,* 30.

[500] Koopman, *Genealogy as Critique*, 93.

[501] Foucault, *Discipline and Punish*, 7.

[502] Foucault, *History of Sexuality, Volume 1: An Introduction (The Will to Know)* (Vintage Books, 1976), 92.

[503] See *Discipline and Punish*, 110.

[504] Foucault, *History of Sexuality*, *Volume 1,* 92.

totalizing force (geographically or temporally), nor an essential form.[505] Hence, rather than advancing a theory of power, Foucault offers an analytic of power, which makes sense of how power operates, circulates, and becomes expressed in particular practices and relations.

In *The History of Sexuality, Volume 1*, Foucault takes as his target a negative account of power, which he refers to as the "repressive hypothesis."[506] According to this received view, power functions primarily as a force of constraint and domination, with truth as its liberatory correlate. Speaking the truth of one's sexuality thus promises emancipation from the constraints imposed by power since repression operates as a ban on truth and thus as an obstacle to knowledge itself. In this framework, power suppresses desire and engenders false consciousness. Speaking the truth of one's sexuality becomes a liberatory act. Similarly, Foucault's *Discipline and Punish* unsettles the assumed inverse relation between power and Enlightenment discourse. Penal reformers have represented the movement from torture to imprisonment as evidence of a diminution of power and the rise of more humane practices. Yet, as Foucault demonstrates, what diminished was not power *tout court* but a particular configuration of sovereign power only to be replaced by disciplinary power.

Crucially, Foucault treats the repressive hypothesis in *The History of Sexuality, Volume 1* and the reduction of penal severity in *Discipline and Punish* not as illusions that obscure truth, but as elements within the very interplay of truth and power that genealogy seeks to diagnose.[507] He does so through three forms of inquiry. First, as a historical investigation, genealogy asks whether the received view of power is really an established fact. For instance, does the disappearance of punishment as public spectacle mark a triumph of reason and progress? Second, as a historico-critical investigation, it asks whether power merely constrains pre-existing objects or is instead productive of new ones. Although it may seem the decrease in corporeal punishment has increased our freedoms, new techniques have produced forms of subjectivity we are largely unaware of. Third, as a historico-political investigation, genealogy asks whether the critical discourses of truth truly contest power or whether they operate within the same historical configuration they denounce. Is it the case that the critical discourse that condemns punishment

[505] Foucault, *History of Sexuality*, *Volume 1*, 94.

[506] Foucault, *History of Sexuality, Volume 1*, Part II

[507] See Dreyfus and Rabinow, *Michel Foucault*, 131.

as a spectacle and celebrates humane treatment really a repudiation of power or is this Enlightened discourse somehow enmeshed with the disciplines themselves?[508] The aim of these questions, Foucault insists, is not to show the received view of power is false, but to integrate it into the general economy from which it emerged.[509]

6. Value Judgements and Explanatory Choice

Having set out the basic elements of a genealogical investigation and determined that it is preferable to archaeology, I now turn my attention to a more direct assessment of the explanatory force of Foucault's genealogy in comparison to Williams's. To do so, I draw on the four criteria I introduced in the previous chapter: determinacy, impartiality, contingency, and regularity. Because it is relatively uncontroversial that Foucault writes about events that have occurred in the determinate past, I will primarily focus on the three criteria of contingency, regularity, and impartiality.

Much as there is little doubt that Foucault's genealogies are situated in determinate time and place, there is likewise little doubt that they are empirically rich. However, I want to resist the temptation to simply posit that more empirical detail entails a better explanation. What makes Foucault's genealogical method preferable to Williams's, then, is not merely its empirical richness, but the values Foucault draws on to aid his selection and representation of significant facts. More specifically, I argue that Foucault's genealogy integrates empirical detail in a manner that satisfies the additional value of complexity of interaction with respect to the explanatory demands of regularity and contingency.

Contingency and regularity are twin criteria. The requirement of contingency places pressure on an explanatory framework to specify a mechanism by which practices become operationalized in a manner that is neither arbitrary nor accidental, but situated against a background of rules. There is, however, a further sense in which the two criteria implicate one another. Just as contingency brings forth the need for regularity, so too does regularity generate the explanatory need for contingency, lest its rules be cast in terms of an underlying law of

[508] Foucault, *History of Sexuality, Volume 1*, 10.

[509] Foucault, *History of Sexuality*, *Volume 1*, 11.

human nature or linguistic practices. What the value of complexity of interaction allows us to assess, I argue, is the extent to which the two constitute each other.

A mechanism is the component of regularity that specifies how individuals adopt particular rules and dispositions. Consider, for example, the mechanism Williams describes in *Truth and Truthfulness*, where the regulation of social dispositions is explained primarily in term of individual and group psychology, such as praise, shame, and blame. Now contrast this with Foucault's account of the mechanisms of disciplinary power in *Discipline and Punish*, which includes techniques such as hierarchical observation, normalizing judgement, and the examination, together with strategies such as documentation, routinized exercises, and architectural arrangements. Although Williams's account is not idealized in the sense of attributing to human agents motivations they do not in fact possess, his minimal account nevertheless comes at a cost. By focusing on psychological motivations, Williams's genealogy omits an explanation of how strategies interact with and shape dispositions in ways that cannot be explanatorily reduced to mental states alone. In this respect, it seems inadequate to characterize the adoption of dispositions as simply the result of other motivations.

For example, I can explain Margaret's motivation to not cheat on an exam partially in terms of her motivation to not be caught cheating, but this explanation would leave out a great deal of ways in which both those motivations are cultivated and operationalized. This is especially important when we take seriously the idea that there are implicit rules that have normative force over action and that operate without our full awareness. It feels more accurate to our everyday experiences that there are factors beyond mental states orienting us towards adopting certain practices. Margaret's disposition to not cheat on an exam, for instance, is operationalized not merely through her fear of being caught, but through an entire architectural and institutional arrangements that cultivate certain dispositions largely independent of her awareness.

In addition to specifying an explanatory "how" by appealing to a mechanism, the criterion of regularity also explains why something occurs by locating it within the broader functional role of the explanandum. As I discussed in the previous chapter, Williams explains our dispositions toward the virtues of truth largely in terms of their function in satisfying basic human needs. Although his genealogy situates these dispositions within a broader historical narrative that extends beyond basic needs, functional needs satisfaction nonetheless takes explanatory priority in his account. Foucault, by contrast, offers a more complex account of mutual interaction

through what he calls the *dispositif*, which is the network that produces needs and practices in the first place. On Foucault's view, functions remain contingent precisely because they are co-produced within a broader strategic network rather than presupposed by an antecedent social need.

Foucault defines a *dispositif* as a "thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions—in short, the said as much as the unsaid."[510] The *dispositif* holds these heterogeneous elements together, much as the *episteme* does in Foucault's archaeological method. Indeed, Foucault characterizes the *dispositif* as a more general form of the *episteme,* encompassing both discursive and nondiscursive formations.[511] In addition to its heterogeneity, Foucault locates the unity of the *dispositif* in its genesis, which involves two moments. First, it arises in response to an "*urgent need*" or strategic objective.[512] Second, it persists through a dual process of "*functional overdetermination*" and "*strategic elaboration*."[513] In other words, a *dispositif* produces effects that exceed the immediate need from which it emerged, while also generating new, unintended outcomes unrelated to the original problematic.

Although Foucault often employs functional language, we should not therefore conclude that a *dispositif* exists to fulfill some preexisting social need or purpose. As he emphasizes, the panoptic *dispositif* "is not simply a hinge, a point of exchange between a mechanism of power and a function; it is a way of making power relations function in a function, and of making a function function through these power relations."[514] Put differently, it is not that power mechanisms emerge in response to an antecedent or universal need, but rather the *dispositif* itself produces the very need the mechanism appears to solve. The *dispositif* conditions the intelligibility of the need as a need and, in turn, makes a particular configuration of power appear as its necessary solution. In this regard, Foucault avoids the pitfalls of functionalist explanations

---

[510] Foucault, "The Confession of the Flesh," in *Power/Knowledge: Selected Interviews and Other Writings, 1972-1997,* ed. Colin Gordon (Vintage Books, 1980), 194-195.

[511] Foucault, "The Confession of the Flesh," 197-198.

[512] Foucault, "The Confession of the Flesh," 195.

[513] Foucault, "The Confession of the Flesh," 195.

[514] Foucault, *Discipline and Punish*, 207.

that, as he cautions in "Nietzsche, Genealogy, History," account for emergence "by appeal to its final term."[515] Instead, he articulates functionalism as part of his historico-critical inquiry, whereby genealogy traces the emergence of an "urgent need" as part of the same configuration as the power mechanisms that come to address it.

The *dispositif* of disciplinary power illustrates this point. According to Foucault, the disciplinary *dispositif* was strategically functional since it emerged in response to the need to "strengthen the social force—to increase production, to develop the economy, spread education, raise the level of public morality," as well as to "improve the exercise of power by making it lighter, more rapid, more effective."[516] The disciplinary *dispositif* continues to persist through a process of functional overdetermination since its formation is tied to broader historical processes involving economic, juridico-political, and scientific structures.[517] Likewise, we can locate one of its strategic elaborations in the state's appropriation and control of disciplinary mechanisms. Although the disciplinary *dispositif* cannot be reduced to the state apparatus, nor can we claim it emerged directly from it, Foucault observes that, through a series of historical elaborations, the disciplines ultimately became coextensive with the police apparatus itself.[518]

Finally, in addition to preferring Foucault's genealogy because its explanation suffices the criteria of contingency and regularity, I suggest that his genealogy is also better at maintaining impartiality with respect to its audience. At the very least Foucault's impartial status is less ambiguous, as it is in Williams's vindicatory genealogy. This is largely because Foucault maintains a strong distinction between descriptive explanation and normative justifications. Nevertheless, this description of Foucault's genealogies is less settled in the literature. Though a more extensive discussion of the relation between genealogy and normative commitment is required, I will conclude by outlining some of the central tensions in this debate.

The claim that genealogies yield normative commitments, as recently presented in Daniele Lorenzini's book *The Force of Truth*, is an outgrowth of the broader normativity debate in

[515] Foucault, "Nietzsche, Genealogy, History," 376.

[516] Foucault, *Discipline and Punish*, 208-209.

[517] See Foucault, *Discipline and Punish*, 218-227.

[518] Foucault, *Discipline and Punish*, 215.

Foucauldian genealogy.[519] Depending on the approach, genealogy can take on subversive, vindicatory, or problematizing dimensions; or, when normatively confused, an element of all three.[520] Lorenzini's argument calls for us to move beyond the vindicatory-subversive dichotomy by recognizing the "possibilizing" dimension of Foucault's genealogies. Understood as such, Foucauldian genealogy possesses normative force by constituting "a concrete framework for action (a transhistorical "we") allowing *genealogy itself* to answer the question 'Why resist?' by generating a sense of ethico-political commitments in its audience."[521] Lorenzini's description here bears resemblance to Queloz's claim that pragmatic genealogies are "justification-yielding," insofar as learning about the emergence and development of practices can affect the space of reasons and "contribute to the vindication or subversion of their product even when that product was not already justified in terms of its formation."[522] Though there are important differences between these two accounts, they converge on the idea that a formation process can itself generate normative commitment, whether for the practical normativity of a conceptual practice or for the normativity of a political struggle.

My concern with both accounts is that a genealogy that aims to yield justifications, even if only as a framework for action, ceases to be audience impartial. As I argued in the previous chapter, what an audience prefers to hear, or is already committed to, should not exert undue pressure on the explanation itself. An explanation does not construct its narrative to suit its audience. It recounts historical truth as accurately as possible. Although it is inevitable that one writes with an audience in mind, the point of the explanation is to truthfully understand something not yet well understood rather than to generate normative commitment. In my view, even when Foucault's genealogies inform justifications, they are descriptive projects before they are normative. They do not aim to convince an audience or tailor a story to its commitments, but

[519] See Nancy Fraser, "Foucault on Modern Power: Empirical Insights and Normative Confusions," *PRAXIS International* 37, no. 1 (1981), 272-87; Amy Allen, *The Politics of Ourselves*, (Columbia University Press, 2008); Koopman, *Genealogy as Critique*; Verena Erlenbusch-Anderson, *Genealogies of Terrorism* (Columbia University Press, 2018); and Amia Srinivasan, "Genealogy, Epistemology, and Worldmaking," *Proceedings of the Aristotelian Society* 119 (2019), 127-56.

[520] See Fraser, "Foucault on Modern Power: Empirical Insights and Normative Confusions."

[521] Daniele Lorenzini, *The Force of Truth: Critique, Genealogy, and Truth-Telling in Michel Foucault* (University of Chicago Press, 2023), 109.

[522] Matthieu Queloz, *The Practical Origins of Ideas* (Oxford University Press, 2021), 215.

to provide the most adequate explanation of how power and knowledge operate in our present moment.

In a 1977 interview, Lucette Finas asks Foucault whether he thinks his writings on the hystericization and psychiatricization of the woman's body in *History of Sexuality* would be of any political help to feminist struggles, to which Foucault responds, "There are a few ideas there, but only hesitant ones, not yet fixed… But it is not up to me to lay down how the book should be used."[523] Like traditional historical inquiry, Foucault's genealogies are primarily descriptive and subsequently remain agnostic regarding the normative conclusions such inquiries will generate. The difference, however, is that Foucault explicitly characterizes his inquiries as practical critique of our present forms of subjectivity. Critique here is not simply contrived to subvert or reject our limits, but to furnish the materials for transgressing those limits, as a "non-positive affirmation."[524] For Foucault, determining what exists beyond that limit, what comes after the act of transgression, is not a task meant for genealogy but for the concrete, experimental practices that genealogy opens up. Nevertheless, one might measure the normative value of a genealogy by its capacity to open new possibilities for subjectivity, while also recognizing that, as with historical research generally, its descriptive value cannot (and perhaps should not) be reduced simply to its practical effects.[525] Even then, what we do with that explanation and how we put it to use is an important consideration, but not one that genealogy is directly involved in.

7. *Fabricating Reality*

In this chapter I have sought to describe the ways in which Foucault draws on and critically reworks some of the central elements of traditional historiography. I have also demonstrated how Foucault's accounts are exemplary of the values involved in a politically realist inquiry, particularly that of empirical adequacy, freedom, and complexity of interaction. Foucault's genealogy, I have argued, is best suited for a politically realist inquiry for two reasons. First, it yields greater explanatory power with respect to our practices under realistic conditions.

---

[523] Michel Foucault, "The History of Sexuality," in *Power/Knowledge*, 192.

[524] Michel Foucault, "A Preface to Transgression," in *Essential Works of Foucault, Volume 3*, ed. James D. Faubion (New Press: 1963/1998).

[525] For an argument in defense of valuing genealogy on the basis of its practical effects, see Bernard Harcourt, "On Critical Genealogy," *Contemporary Political Theory* 24 no. 2 (2025), 167-185.

Second, it provides critical insight to our practices in ways that the other explanatory methods I have discussed cannot. Political realists have much to gain from Foucault's critical approach to political reality and the possibilities he leaves open for freedom.

In that same 1977 interview, Lucette Finas asks Foucault whether his genealogies are fictions, to which he responses, "I have never written anything but fictions. I do not mean to say, however, that truth is therefore absent. It seems to me that the possibility exists for fiction to function in truth, for a fictional discourse to induce effects of truth, and for bringing it about that a true discourse engenders or 'manufactures' something that does not as yet exist, that is, 'fictions' it. One 'fictions' history on the basis of a political reality that makes it true, one 'fictions' a politics not yet in existence on the basis of historical truth."[526] Fiction here is entangled with truth. Indeed the "fiction" of history begins within political reality precisely because that reality has been itself fabricated. The effects of power are fictitious insofar as we have made them up. This does not make them any less real. What the effects of power are not, importantly, are illusory or false. Foucault's genealogies are thus attempts to construct a narrative from these political realities, and in turn, to enable a fictional (because not yet realized) politics on the basis of historical truth. What he seeks to avoid is an explanation that uncritically accepts our present as reality or one that posits reality on the basis of an uncritical fiction.

Thus far in my dissertation, I have remained at a high level of methodological abstraction, first by specifying what a politically realist justification should look like, both in its aims and its procedures, and then by identifying the criteria appropriate to politically realist inquiry in debates concerning genealogical explanation. I now turn to mobilizing this methodological discussion of dignity and its constitutive practices, a value I argue stands in need of historical and philosophical clarification. I call this method a *politically realist genealogy*, and it is to this explanation that I now turn.

[526] Foucault, "The History of Sexuality," in *Power/Knowledge*, 193.

VI.

A POLITICALLY REALIST GENEALOGY OF DIGNITY

"Judicial opinions in this area do not make pleasant reading."
—Justice William Brennan, *Rhodes v. Chapman*

On September 9, 1971, more than 1,000 prisons took control of Attica Correctional Facility, seizing 39 guards as hostages in the process. After four days of futile negotiations, Governor Nelson Rockefeller ordered for the prison to be retaken. At 9:46 am, pursuant to Rockefeller's executive order, the New York State Police broadcasted over the radio for the militia to move into Attica's D Yard. At 10:05 am, the NYSP had declared Attica fully secured.[527] In less than half an hour, 128 men had been shot, 43 people had been killed of which 33 of them were prisoners and 10 prison staff. On this gruesome day, the Attica prison uprising marked the highest number of fatalities in the history of U.S. prison uprisings.

Although the uprising began spontaneously, it would be inaccurate to say that the prison administrators and staff were surprised by the event. Tensions had been rising at Attica for quite some time. In September 1971, Attica prison housed over 2,200 people in a facility meant for 1,600. The prison cells were poorly ventilated and inadequately heated, the work was arduous with few prisoners earning more than six cents per day, the conditions were unhygienic as prisoners were only allotted one roll of toilet paper and one bar of soap per month, arbitrary punishment for trivial infractions was rampant, and the men were starving.[528]

In June 1971, the Attica Liberation Faction was founded. By July, its members had presented a list of 27 demands to the Commissioner of Corrections, such as an end to physical brutality, segregation and political persecution, arbitrary punishment, and unsanitary eating conditions; alongside positive requests, such as more food and water, a set of clear standards for prison

[527] Heather Ann Thompson, *Blood in the Water* (Pantheon Books, 2016), 187.

[528] Thompson, *Blood in the Water*, 11.

officials to follow, an oversight committee to evaluate prison practices, and basic legal assistance. The group concluded the list of their demands stating:

> We are firm in our resolve and we demand, as human beings, the dignity and justice that is due to us by our right of birth. We do not know how the present system of brutality and dehumanization and injustice has been allowed to be perpetrated in this day of enlightenment, but we are the living proof of its existence and we cannot allow it to continue. The taxpayers who just happen to be our mothers, fathers, sisters, brothers, daughters and sons should be made aware of how their tax dollars are being spent to deny their sons, brothers, fathers and uncles of justice, equality, and dignity.[529]

We might pause here to consider why demands for justice and equality alone were not sufficient, and what a claim to dignity adds that the other two do not already contain. An answer to this question involves giving a broader account of dignity, one in which the dignity of the prisoner first appears in 1958, reemerges in 1966, and continues up until the present moment that is infused with demands for dignity, or at the very least a reflex against indignity. Such an account would reveal the Attica Liberation Faction's demand for dignity was not accidental, but part of a wider deployment of the concept in United States prisons.

By most accounts the Attica Prison Riot was a failure. Following nearly fifty years of stable incarceration rates, the prison population began to drastically rise in the mid-1970s. When Ronald Regan took office in 1980 the total prison population was 329,000, already a large jump from nearly a decade prior.[530] When he left office eight years later, the prison population had grown to 627,000.[531] With the rise of mass incarceration came a turn away from the rehabilitative hope of the early 1970s toward a deeply retributive, "tough-on-crime" approach to punishment. Much like the violent snuffing of the Attica prison uprising, the prisoners' rights movement was largely suppressed within a few years of its emergence. Rather than viewing the prisoner rights movement as a failure, however, I suggest we consider the kinds of practices it produced. What undoubtedly emerged from this period and what continues today is the frequency with which we use dignity in politics, in law, and in ethics. But the popularity of the term and the ease with

---

[529] Attica Liberation Faction, "Manifesto: The Attica Liberation Faction Manifesto of Demands and Anti-Depression Platform, 1971" *Black Agenda Report,* effective April 13, https://blackagendareport.com/manifesto-attica-liberation-faction-manifesto-demands-and-anti-depression-platform-1971

[530] U.S. Justice Department, Office of Justice Programs, Bureau of Justice Statistics. *Prisoners in 1980* (Washinton, DC, 1981), 2-3.

[531] U.S. Justice Department, Office of Justice Programs, Bureau of Justice Statistics. *Prisoners in 1988* (Washinton, DC, 1989), 2-3: 2-13.

which we intuitively grasp its meaning has made the concept unwieldy and in need of greater specification.

There are many approaches one could take to clarify dignity's point and value in our contemporary moment. The approach I take is genealogical in design. To make sense of dignity's emergence, its point and value, and how it becomes operationalized, I develop a *politically realist genealogy*, which is a method that describes the emergence and development of thick ethical concepts. Given that human dignity is viewed as universal, my genealogy thus aims to reveal it is, in fact, historically contingent. Guiding my inquiry is the metaethical view that our values are thoroughly pluralistic, which requires me to narrow my inquiry to a localized instantiation of dignity. In my case, I focus on dignity practices in United States prisons, an area where dignity has proven to be especially thick, as it were. I first clarify the descriptive content of dignity by specifying the conditions under which it emerged as a value in punishment jurisprudence. I then move to the evaluative content of dignity, by situating it within a network of other thick ethical and counter ethical concepts that frequently accompany dignity in legal discourse. Finally, I consider the standards that render dignity intelligible in prisons, alongside the procedures by which compliance with those standards is evaluated.

When Foucault wrote of the emergence of the prison and of disciplinary power, he primarily focused on the delinquent as one of power's products. My genealogy shifts his focus to examine the prison administrator. This allows us to better understand how legal rulings, dispositions, techniques, and instruments generated standards expressive of "dignified treatment," standards that are themselves objects of knowledge and sites of power. My genealogy also aims to specify the concrete practices and affiliated values that construct the meaning of dignity in United States prisons. As Attica Liberation Faction's demands reveal, I argue that dignity does not function to express the highest moral treatment owed to individuals in virtue of their status or membership. Rather, it functions as a threshold concept that structures the two poles of legalism. On the one hand, dignity curtails the overly strict and mechanistic application of rules and ensures the individual is seen and treated as an individual; on the other hand, it establishes clear standards for and constraints on prison administrators and staff in their treatment of prisoners. To make a demand in the name of dignity, on my view, is to identify forms of treatment for which there are no penological justifications.

1. Politically Realist Genealogies

As I mentioned in the previous chapter, one of the defining features of Herodotus's *Histories* is a first-person reflection on historical practice, often described as a *meta-history*. The purpose of this reflection is to set expectations and orient the audience, particularly when a method departs from established approaches. With this in mind, I will briefly characterize the defining elements of a politically realist genealogy before demonstrating this method through the concept of dignity.

As I have demonstrated in the previous chapters, there is no single way of carrying out a genealogical method. Genealogical projects typically draw on a range of philosophical influences rather than a unified methodological template. Williams, for instance, combines methodological elements found in both Hume and Nietzsche, the latter of whom also serves as a central source of inspiration for Foucault. In turn, these figures have shaped subsequent genealogical work, both at the level of metaphilosophical reflection on genealogy and in its concrete deployment in historical and political analysis.[532]

With this in mind, I begin my politically realist genealogy (hereafter referred to as simply a genealogy), by focusing on an ethical concept we take to be obvious, or obviously without history and demonstrate that this ostensibly self-evident concept has, in fact, a precise historical emergence and development. A genealogy is a narrative that explains ethical practices by describing the way in which it came about. For example, a genealogical explanation uncovers the historical conditions in which the ethical practice of dignity became an intrinsic value and, in turn, a principle that guides ethical action. Put differently, genealogy makes sense of the conditions under which dignity became a reason for action and then traces the concrete practices of that action. "Making sense" here is a term of art drawn from Williams and is largely an interpretive technique for rendering an action intelligible from an internal point of view. When an explanation makes sense of someone's action, it does so in terms of her reasons, motivations, and

[532] For example, Verena Erlenbusch-Anderson, *Genealogies of Terrorism: Revolution, State Violence, Empire* (Columbia University Press, 2018); Miranda Fricker, "The Genealogy of Testimonial Injustice," in *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford University Press, 2007);); Colin Koopman, *Genealogy as Critique: Foucault and the Problems of Modernity* (Indiana University Press, 2013)*; How We Became Our Data: A Genealogy of the Informational Person* (University of Chicago Press, 2019); Daniele Lorenzini, *The Force of Truth: Critique, Genealogy, and Truth-Telling in Michel Foucault* (University of Chicago Press, 2023); *Ladelle McWhorter, Racism and Sexual Oppression in Anglo-America: A Genealogy* (Indiana Press University, 2009); *Unbecoming Persons: The Rise and Demise of the Modern Moral Self* (University of Chicago Press, 2025); Matthieu Queloz, *The Practical Origins of Ideas* (Oxford University Press, 2021).

commitments. Genealogy thus makes sense of an ethical concept by explaining how it became a reason for action in a situation of practical deliberation.

By making sense of an ethical concept in historical terms, a genealogy constructs a thick ethical concept. As I discussed in Chapter Two, a thick ethical concept is both action-guiding, insofar as it is evaluative or prescriptive, and also guided by the world, insofar as it is descriptive. In other words, a thick ethical concept guides action precisely because it conjures an image of that concept that has been shaped by the world. In this sense, someone can correctly or incorrectly apply a thick ethical concept. In order to make sense of a thick ethical concept it is important to situate it alongside other ethical concepts and importantly, counter-ethical concepts. This will help provide a fuller range of the motivations that hold a single concept together.

It does not follow that all individuals will unanimously agree on a thick ethical concept's correct application. Indeed, in many cases agreement comes easily while in others there will be considerable disagreement.[533] Part of what my genealogy is concerned with then, is understanding the conditions under which agreement is reached and as such, when concepts indeed guide action. In such cases, we might say that the concept is thought by agents to have intrinsic value, though as I clarified in Chapter Four, intrinsic value here is meant to convey something like ethical importance. When a concept is viewed as having intrinsic value, it means that I give it deliberative priority in my considerations. When I give a concept deliberative priority, it means that when the situation arises, I have good reasons for it to guide my action. It is my view that when we agree on a correct application of an ethical concept, that is, when we agree that it should guide action in a particular situation, we have made sense of that concept as having value, and as being important in justifying our actions. In cases where disagreement abounds, it may be the case that it is unclear whether a concept figures in guiding action or that it may be ranked lower in deliberative priority among other concepts. In such a situation, the concept would no longer hold intrinsic value.

In the political realism I have discussed, particularly in Shklar and Williams, there exists a claim for other realists to recognize the importance of ethical life within politics, whether this is because one sees the ramifications of politics from the perspective of one of its victims, as Shklar did, or because one sees ethical life as an important dimension of our human condition (if only among other important dimensions), as we find in Williams. In either case, the point is that

[533] Williams, *Ethics and the Limits of Philosophy* (Harvard University Press, 1985), 141.

political realism need not simply be aimed at solely pointing out the flaws in traditional political philosophy, but can itself provide normative guidance, albeit one grounded in, as Williams notes, committed beliefs supported by evidence and clear-headed desires articulated with reasons.[534] To arrive at such normative guidance, realists argue we require conscious reflections on our present political circumstances, which makes it a natural compatriot of genealogy.

It is by now a platitude to observe that Foucault's genealogies are "histories of the present," yet this insight remains central insofar as a genealogy is oriented towards the problems we presently face. In this sense, genealogy supplies a method for conscious reflection on the ethical concepts so thoroughly embedded in contemporary discourse and brings them into critical view. And once a genealogy has constructed a thick ethical concept by making sense of the conditions under which it guides actions or fails to guide action, political realism can construct normative principles that are historically grounded, morally austere, but nevertheless politically ambitious.

In the genealogical explanation that follows, I will offer a deliberately minimalist narrative in the sense that a great deal will be left out. This is not, in itself, a defect, though it does introduce certain dangers and difficulties that are worth flagging at the outset. One such danger lies not only in what a minimalist narrative omits, but what it fills in. In particular, my story risks overestimating the extent to which historical actors acted in service of dignity in a way that is far more rational than it actually was.[535] One way a minimalist genealogy can guard against over-rationalization is by foregrounding contingency and by acknowledging that events could have unfolded in another way than my story suggests, and that there was a great deal more elements guiding action than simply the ethical concept under consideration. Methodologically, this requires treating motivational states not as the primary source of action, but as effects of events and specific historical configurations that we can nevertheless make sense of as having guided action. A genealogy of an ethical concept is therefore not an investigation into the essence of that concept or an analysis of its necessary and sufficient conditions, but rather a historical investigation into the events that have constructed such a concept into a reason for action. For instance, it is a fact that the Greek concept "*kleos*" once led men into battle. We can acknowledge this without reducing those actions simply to *kleos*, all while recognizing that *kleos*, as a

[534] Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton University Press, 2002), 197.

[535] See Williams, *Truth and Truthfulness*, 248.

motivational state, is itself an effect of historical events that a historical account cannot fully rationalize. Notwithstanding these problems, I believe an awareness of reductive explanations can go a long way toward helping us avoid them. Nor should these methodological difficulties stop us from engaging in genealogy. To put it in Williams terms, "[t]he past will not make sense unless we make sense of it."[536]

Dignity is ubiquitous in moral and political philosophy. In discussions of care, trust, respect, and equality, dignity often serves as the ultimate justification for our moral relations. Yet the dignity literature is of no avail in producing clarity or consensus. Efforts to specify dignity's normative grounds and content, or to lend it a precise definition only generate more factions among dignitarians and leave the reader feeling overwhelmed and confused.

In his essay, "Dignity's Gauntlet," Remy Debes observes a customary habit in dignity discourse. When introducing dignity, one first traces its historical-etymological origins and then notes the dominance of Kantian and Judeo-Christian traditions that link our present concepts to rational autonomy or a spiritual identity with God, respectively.[537] From here, one rehearses the debates over dignity's grounds and its defining properties, before considering whether dignity is unique to humans or whether it can be extended to a wider moral community. At the risk of running the same course, I introduce a similar set of distinctions intended to motivate the need for a metaethical approach that treats dignity as it presently exists in localized contexts rather than as part of a substantive moral theory, which specifies how dignity ought to be.

We can distinguish two broad ways of thinking about dignity, both of which can be at work within a given theory. First, there are accounts concerned with the normatively grounding content of dignity. Martha Nussbaum grounds dignity in inherent worth, which is then realized through her ten central capabilities.[538] Jürgen Habermas and Seyla Benhabib ground dignity in the communicative freedom of the person.[539] George Kateb locates dignity in the distinctness of

---

[536] Williams, *Truth and Truthfulness*, 244.

[537] Remy Debes, "Dignity's Gauntlet," *Philosophical Perspectives* 23, no. 1 (2009), 45.

[538] Martha Nussbaum, "The Worth of Human Dignity: Two Tensions in Stoic Cosmopolitanism," in *Philosophy and Power in the Graeco-Roman World: Essays in Honor of Miriam Griffin*, ed. Gilian Clark and Tessa Rajak (Oxford University Press, 2002), 31-50; Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Belknap Press, 2007); Nussbaum, *Creating Capabilities* (Harvard University Press, 2011).

[539] Seyla Benhabib, *Dignity in Adversity: Human Rights in Troubled Times* (Polity, 2011); Jürgen Habermas, "The Concept of Human Dignity and the Realistic Utopia of Human Rights," *Metaphilosophy* 41, no. 4 (2010), 464-480.

the human species apart from and superiority to all other earthly creatures. He writes, "The core idea of human dignity is that on earth, humanity is the greatest type of being."[540] Will Kymlicka challenges Kateb's view and argues that dignity need not rely on human supremacism. Instead, we can ground it in a broader moral community that includes both humans and nonhumans.[541] In a manner consistent with Kymlicka's view, feminist care ethicists ground dignity in the concrete particularity of individual persons.[542]

Second, there are theories concerning the normatively expressive content of dignity. These theories are less focused on identifying dignity's grounds or properties and more on what kinds of moral or legal claims the concept expresses. Such accounts maintain that dignity is constructed rather than inherent. On Stephen Darwall's view, dignity expresses the value or worth of person and is constructed through practical reason, which yields unconditional and authoritative claims made on us. Human dignity, for Darwall, is tied to the second-personal character of our relations, including the authority to demand certain forms of treatment from one another.[543] For Jeremy Waldron, dignity is a moral and legal principle that expresses rank or status. What once was a status reserved for the few has now been democratized and extended to all humans. Dignity now expresses the "idea of the high and equal rank of every human person" and is "intimately connected with the idea of rights—as the ground of rights."[544] On this view, dignity is a result of a historical, legal, and moral transformation evincing the egalitarian valuation of all persons. This helps explain how the term can both describe a status we all have and "convey the demand that rank or status should be recognized."[545] The fact that dignity is something to be protected points to its fragility. Adam Etinson, who shares with Waldron that

---

[540] George Kateb, *Human Dignity* (Harvard University Press, 2001), 3.

[541] Will Kymlicka, "Human Rights Without Human Supremacism," *Canadian Journal of Philosophy* 48, no. 6 (2018), 763-792.

[542] Robin S. Dillon, "Respect and Care: Toward Moral Integration," *Canadian Journal of Philosophy* 22, no. 1 (1992), 105-131; see also Eva Feder Kittay, "Equality, Dignity, and Disability," in *Perspectives on Equality*, ed. Mary Ann Lyons and Fionnuala Waldron (Liffey Press, 2005), 95-122.

[543] Stephen Darwall, *The Second-Person Standpoint: Morality, Respect and Accountability* (Harvard University Press, 2006), 13; this view is also shared by Anthony Appiah, *The Honor Code: How Moral Revolutions Happen* (Norton Press, 2010).

[544] Jeremey Waldron, *Dignity, Rank, and Rights* (Oxford University Press, 2012), 210.

[545] Waldron, *Dignity, Rank, and Rights*, 17.

dignity expresses status or social standing, argues that the concept requires "that we avoid subjecting people to the specific (socially oriented) harm of humiliation or degradation and, more positively, that we help protect them from such harm too."[546] Dignity thus expresses a claim against certain forms of treatment.[547]

This brief overview only begins to capture the range of positions in the dignity literature, in which, Debes notes, "Dignity has been conceived of as honor, rank, station, inherent worth, inalienable worth, equal worth, supreme worth, uniqueness, beauty, poise, *gravitas*, personality, integrity, bodily integrity, self-respect, self-esteem, a sacred place in the order of things, simply brute and unquestionable "specialness," and even, if she is being exhaustive, the apex of astrological influence."[548] Given this range, is it any surprise that it faces skepticism? A central critique is that dignity's breadth renders it too indeterminate to ground substantive moral or legal claims. Dignitarians have stretched the concept so thin that it often fails to guide the action they purport it ought to. Dignity is presented as the bedrock of ethical life because it is self-evident, yet it repeatedly disappoints us when it does not function as a reason for action. Andrea Sangiovanni acknowledges dignity's intuitive appeal, but argues that no account has yet to sufficiently explain how we are equal in dignity or why we have it.[549] Ruth Macklin goes further and claims that dignity is merely a placeholder for other more substantive terms. Dignity, she writes, "is a useless concept. It means no more than respect for persons or their autonomy."[550] Finally, Samuel Moyn argues that dignity's theological underpinnings hold no promise for liberal secularism and may even weaken the protection of human rights.[551]

In light of these disagreements, a genealogical approach to dignity is especially illuminating. I argue an outright dismissal of dignity by the skeptics reveals a lack of historical curiosity. It

---

[546] Adam Etinson, "What's So Special About Human Dignity?" *Philosophy & Public Affairs* 46, no.4 (2020), 355.

[547] This view is also held by Suzy Killmister, *Contours of Dignity* (Oxford University Press, 2016).

[548] Debes, "Dignity's Gauntlet," 46.

[549] Andrea Sangiovanni, *Humanity Without Dignity* (Harvard University Press, 2017), 15.

[550] Ruth Macklin, "Dignity is a Useless Concept," *BMJ* 326, no. 20-27 (2003), 1418-1420.

[551] Samuel Moyn, *Human Rights and the Uses of History* (Yale University Press, 2014); Moyn, "The Secret History of Constitutional Dignity," *Yale Human Rights & Development Law Journal* 25 (2012); see also Neomi Rao, "Three Concepts of Dignity in Constitutional Law," *Notre Dame Law Review* 86, no. 1 (2013), 183-272.

may be that dignity's apparent thinness is not a defect of the concept itself, but a failure of philosophy to apprehend it as thick, as shaped by the world and therefore temporally and geographically situated. In the genealogy that follows, it is my hope to add some historical substance to this elusive concept.

2. The Birth of Dignity

In an 1826 manual, *A Brief Account of the Construction, Management, and Discipline of the New York State Prison at Auburn*, Gershom Powers, the keeper of Auburn Prison, set forth the principles of the Auburn system, one of the most influential American penal methods of the nineteenth century.[552] Within this manual, the concept of dignity appears, but it carries a sense distinct from the one that will take shape in the latter half of the twentieth century. In this older sense, dignity is closely associated with status and is notably not an attribution of the individual, but one of institutions, professions, and laws. For example, a legislation justly appreciates its own dignity and the character of the laws, sheriffs are required to support the dignity of the laws, and the courts are expected to maintain their dignity.[553] Although this earlier usage does not disappear altogether, it will nevertheless sound outdated when set against its contemporary counterpart.

Dignity, as an attribute of a profession, appears again thirty-four years later, in 1870, when leading officials in American corrections met in Cincinnati, Ohio, to form the American Prisons Association and to articulate a set of principles of punishment. As an appendix to the more routine descriptions of prison administration, the group published the "Declaration of Principles," which it subsequently revised and reaffirmed at the Sixtieth Annual Congress of the American Prison association in 1930, and again in 1954, by which time the organization has assumed its current name of the American Correctional Association. Across all three versions of the document, dignity is again invoked in the "Declaration of Principles," with a sense largely continuous with that found in the Auburn manual and in earlier correctional manuals from 1870 to 1954.

---

[552] Gershom Powers, *A Brief Account of the Construction, Management, and Discipline of the New York State Prison at Auburn* (U.F. Doubleday, 1826).

[553] Gershom Powers, *A Brief Account of the Construction, Management, and Discipline of the New York State Prison at Auburn*, 53-53.

Principle VII, for instance, stresses the importance of specialized training and the cultivation of "high qualities of head and heart," in the formation of a good prison or reformatory officers.[554] Only once prison officials achieve scientific and uniform training methods, the manual asserts, will it be raised "to the dignity of a profession."[555] Dignity here functions as a reason for establishing "schools for the training of prison executive guards" across the United States.[556] Principle XXV invokes a clearly related sense of dignity. Investigations into mental disorder and its relation to crime, the principle specifies, should be "conducted with greater knowledge, dignity, and fairness," thereby ensuring that "the restraint of the insane," is both "certain" and "humane."[557] Once again, dignity is a concept that guides an investigation done thoroughly and with integrity so as to achieve certainty and fairness.

After maintaining relative consistency in its principles for nearly eighty-five years, the revised *Manual of Correctional Standards* of 1966 marks a notable shift in the use of dignity. References to dignity as a matter of professional status largely disappear from the principles, thus making room for the emergence of a new conception of dignity. Alongside a broader revision in principles that place greater emphasis on an attention to the individual offender as opposed to the construction of the profession, Principle XVI provides a clear articulation of this new sense of dignity:

> The principles of humanity and human dignity to which we subscribe, as well as the purposes of rehabilitation require that the offenders, while under the jurisdiction of the law enforcement and correctional agencies, be accorded the generally accepted standards of decent living and decent human relations. Their food, clothing and shelter should not be allowed to fall below the generally accepted standards, and they should be afforded the conventional conveniences made possible by our technological progress. Their health needs—both physical and mental—should be met in accordance with the best medical standards. Recreation should be recognized as a wholesome element of normal life.[558]

---

[554] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1954), 409.

[555] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1954), 409.

[556] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1954), 409.

[557] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1954), 411.

[558] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1966), xxi; Examples indicating this broader shift towards the individual can be find in Principle XIII: "Correctional institutions and agencies can best achieve their goal of rehabilitation by focusing their attention and resources on the complete study and evaluation of the individual offender and by following a program of individualized treatment," and in

In 2018, when the American Corrections Association again revised and pared down the principles, dignity here is afford a more central role:

> HUMANITY. Corrections upholds the dignity of individuals, the rights of all people and the potential for human growth and development. Social order in a democratic society depends upon full recognition of individual worth and respect for the dignity of all its members; therefore, laws, administrative policies and corrections practices must be governed by this principle and measured against standards of fairness and decency, whether applied to those under corrections care and control, its staff, crime victims or the general public.[559]

In these principles, dignity appears alongside other ethical concepts, such as fairness, respect, and decency, and they specify forms of treatment that the administrator affords to the individual qua human being. In particular, the 1966 manual specifies these treatments through standards, which are notably absent in the revised 2018 manual. Nevertheless, to subscribe to human dignity, to use it as a guide for treatment of prisoners, means to establish "generally acceptable standards" of decent living and human relations, particularly as it relates to food, clothing, shelter, and medical needs.

Though dignity as the inherent worth of an individual is largely understood to be timeless, it has a relatively precise emergence, at least in the American prison. How, then, do we make sense of this emergence of a new dignity in the latter half of the twentieth century? What are the kinds of ethical and counter-ethical concepts it appears alongside, and what are the concrete forms of treatment it specifies? Answering these questions, I suggest, will construct dignity as a thick ethical concept in American prisons and will better orient us towards a normative project in which dignity, already efficacious for many forms of treatment, can guide politically ambitious projects.

In making sense of this historical emergence, I will not attempt to offer a causal explanation of what specifically brought dignity about, though I would be remiss not to mention the Universal Declaration of Human Rights (1948), the International Covenant on Civil and Political Rights (1966), the Civil Rights movements of the 1960s, and the global protests of 1968, among

---

Principle XIV: "The sentence or disposition determining the treatment for the offender should be based on a full consideration of the social and personality factors of the particular individual" (xxi).

[559] American Correctional Association, *A Manual of Correctional Standards* (American Correctional Association, 2018).

other many more political movements in which dignity featured. Recognizing this period as one of political disturbance and politically ambitious reforms, it is clear that, as dignity claims saturated international political discourse, the concept gradually made its way into American caselaw and correctional standards. The task of this genealogy is therefore to make sense of the world in which the concept emerged and the action it has guided within that world.

One area in which dignity has featured most prominently as a reason guiding rulings is in American federal Eighth Amendment case law.[560] Although dignity arguably features as an animating principle in all Eighth Amendment case law since *Trop v. Dulles* (1957), in which the Court declared "[t]he basic concept underlying the Eighth Amendment is nothing less than the dignity of man," my analysis remains more textual in scope. For now, I consider only those cases in which dignity is explicitly invoked as a reason for or against a particular mode of treatment in both Supreme Court decisions and lower federal court cases that have shaped decisions in the higher courts.

Within this body of law, I understand dignity claims in American prisons operating in three distinct, though related "dignity types", each of which presupposes dignity as ascribed to the individual. First, proportionality-based dignity specifies a dignity violation when punishment is disproportionate to the offense. Although proportionality cases occupy an important place in Eighth Amendment jurisprudence, explicit invocations of dignity of this type are relatively rare. Indeed, the Supreme Court has decided numerous cases under the Cruel and Unusual Punishment clause addressing proportionality, dignity often does not figure prominently in the Court's reasoning.[561] Second, culpability-based dignity violations arise when punishment is imposed without adequate consideration of the offender's culpability. These cases involve assessments of culpability that take into account mitigating factors, such as facts of the crime or characteristics of the offender, and/or the offender's rational capacities. Dignity claims of this sort became especially prominent in the early 1970s in the context of death penalty jurisprudence and have persisted in subsequent decades. Third, treatment-based dignity violations occur when the

[560] The Eighth Amendment states, "Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishment inflicted" (U.S. Const. amend. VIII).

[561] See for example, *Weems v. United States*, 217 U.S. 349 (1910), which addressed barbarous and torturous punishment for falsifying public documents; *Robinson v. California* 370 U.S. 660 (1962), which struck down punishment of any kind on the grounds of an individual's status, as well as cases concerning capital punishment such as *Coker v. Georgia*, 433 U.S. 584 (1977), *Gregg v. Georgia*, 428 U.S. 153 (1976), *Proffit v. Florida* 428 U.S. 242 (1976), and *Jurek v. Texas* 428 U.S. 262 (1976).

conditions of confinement are so egregious that they exceed any legitimate penological purpose. Dignity functions here as a threshold or minimum standard and marks the line between forms of treatment that, though harsh, are justified by the state as penologically necessary, and those that are considered so severe that they only exist to harm and humiliate.

3. Type I: Proportionality Dignity Violations ca. 1957

In 1944, United States Army private Albert Trop escaped from a military stockade, only to be apprehended by a search party soon after. Trop was subsequently convicted of desertion in a general court martial and was sentenced to three years hard labor, loss of pay, and given a dishonorable discharge. Eight years later, in 1952, when Trop filed an application for a passport, he was surprised to learn that he had lost his citizenship. Due to a clause in the 1940 Nationality act, Trop's wartime desertion had made him stateless. The question put forth by the Court was whether the constitution permits congress to take away citizenship as a punishment for the crime of wartime desertion or whether it violates the Cruel and Unusual Punishment clause. [562]

In its ruling, the Court considered whether the state could revoke someone's citizenship if he had failed to fulfill their basic responsibilities and duties as citizens. The Court began its ruling by noting that the legality of the death penalty should not serve "as an index of the constitutional limit on punishment," nor should its existence license the government "to devise any punishment short of death within the limit of its imagination."[563] Though death might be a distinct form of punishment in kind, since its finality is unmatched once the punishment is meted out, the Court's characterized statelessness as its own kind of death. The stripping of citizenship involves "no physical mistreatment, no primitive torture," but instead, "the total destruction of the individual's status in organized society. It is a form of punishment more primitive than torture, for it destroys for the individual the political existence that was centuries in the development… In short, the expatriate has lost the right to have rights."[564]

[562] *Trop v. Dulles*, 356 U.S. 86, 100 (1958)

[563] *Trop v. Dulles*, 356 U.S. 86, 99 (1958)

[564] *Trop v. Dulles*, 356 U.S. 86, 101-102 (1958)

Losing one's citizenship amounts to a political death and places the offender in a state of great fear and distress.[565] The harm of statelessness, the Court deemed, is not its physicality but its power to erode the psychological capacities of the offender since he is rendered vulnerable to whatever fate might befall him in the future. In losing his right to rights, the offender is forced into a life of constant discrimination and with no recourse to challenge it. The Court argued this amounts to little more than "naked vengeance," that asserts that "he who refuses to act as an American should no longer be an American."[566] Stripping someone of citizenship is to make him an outcast of society and to cut him off from rehabilitation. What is more, the punishment serves no penological purpose of protecting society. Though the offender is stateless, he is not banished from the land. Indeed, his derelict status and his lack of rights is likely to force him into a life of crime, thus placing society at *greater* risk than before.[567]

The harms of stripping someone of his citizenship amounts to a violation of dignity, a line of reasoning the Court adopted from the lower federal courts, which held that the "American concept of man's dignity does not comport with making even those we would punish completely stateless."[568] In its reasoning, the Supreme Court offered a further articulation of dignity, thereby reifying the concept as the animating principle of the Eighth Amendment. In *Trop*, the Court established, "The basic concept underlying the Eighth Amendment is nothing less than the dignity of man. While the State has the power to punish, the Amendment stands to assure that this power be exercised within the limits of civilized standards."[569] Further, the Court acknowledged that prior to *Trop*, it "has had little occasion to give precise content to the Eighth Amendment," and noted that this was in part due to the fact that the words of the Amendment are not static, but draw their meaning "from the evolving standards of decency that mark the progress of a maturing society."[570] So then, a core relation set forth here between punishments

[565] *Trop v. Dulles*, 356 U.S. 86, 102 (1958)

[566] *Trop v. Dulles*, 356 U.S. 86, 112 (1958)

[567] *Trop v. Dulles*, 356 U.S. 86, 111 (1958)

[568] *Trop v. Dulles*, 356 U.S. 86, 101 (1958)

[569] *Trop v. Dulles*, 356 U.S. 86, 100 (1958)

[570] *Trop v. Dulles*, 356 U.S. 86, 100-101 (1958); see also *Weems v. United States*, 217 U.S. 349 (1910), for the Court's reasoning.

that violate the Cruel and Unusual Punishment clause because they violate the dignity of man, are largely dependent on historical interpretation rather than on some predefined notion of a dignified life.

This articulation of dignity in *Trop* formed the citational bedrock of subsequent rulings. Yet, despite the number of cases addressing the proportionality of punishment under the Cruel and Unusual Punishment clause, dignity, as an explicit reason against disproportionate punishment, largely disappears from the case law and reemerges only much later in *Kennedy v. Louisiana*. In that case, a Louisiana man, Patrick Kennedy, was found guilty of raping his eight-year-old stepdaughter and sentenced by a jury to death. The question put before the Court was whether the death penalty for the crime of child rape was a violation of the Cruel and Unusual Punishment clause.

The Court began its reasoning by declaring that, "Evolving standards of decency must embrace and express respect for the dignity of the person, and the punishment of criminals must conform to that rule," part of which requires a penological justification for punishment, especially when death is involved.[571] Unlike *Trop*, which primarily considered whether the disproportionality of the punishment was so cruel as to amount to a violation of dignity, *Kennedy* considered whether the punishment was so *unusual* as to amount to a violation of dignity. The Court focused its reasoning around the idea of social consensus, comprised of contemporary norms, including how frequency with which death penalty sentences and executions are meted out for nonhomicide crimes.[572]

As the Court noted, even though nine states had statutes permitting capital punishment for adult or child rape, no individuals had actually been executed for rape since 1964 nor had prisoners been executed for any other nonhomicide crime since 1963.[573] Indeed, it noted that, at the time of the case, Louisiana had been the only state since 1964 to sentence someone to death for the crime of child rape and only two individuals were on death row for nonhomicide crimes. What this indicated was a reluctance on the part of state legislation and juries to permit the execution of prisoners for nonhomicide crimes, however contemptible those crimes may be.

---

[571] *Kennedy v. Louisiana*, 554 U.S. 407, 420 (2008)

[572] *Kennedy v. Louisiana*, 554 U.S. 407, 434 (2008)

[573] *Kennedy v. Louisiana*, 554 U.S. 407, 433 (2008)

The Court reasoned that pursuing the death penalty for a nonhomicide crime, such as rape, faced serious challenges. This is because there was a lack of meaningful standards to guide juries in separating out the "most severe cases of child rape" from ones that were less severe.[574] Even if aggravating factors were compared against mitigating factors in the meting out of punishment, the Court suggested that the crime "will overwhelm a decent person's judgement" and subsequently erode our confidence in the jury's ability to reach a decision that was not "so arbitrary as to be "freakis[h]," *Furman, supra*, at 310 (Stewart, J., concurring)."[575] This, when paired with how rare the death penalty was at the time for such a crime due to social consensus on its disproportionate nature and its lack of penological justification, suggested that a death penalty sentence for this type of crime is *de facto* arbitrary. As such, the Court ruled that the Eighth Amendment barred states from imposing the death penalty for the rape of a child, where the crime did not result, and was not intended to result, in the child's death.

In sum, the crimes and punishments at issue in *Trop v. Dulles* and *Kennedy v. Louisiana* are distinct. What the two cases share is that their punishments were so severe *in* kind relative to the crime committed as to have violated the offender's dignity. In both cases, the punishments were found to serve no penological purpose that a lesser punishment would not equally accomplish. In *Trop*, the punishment served only to seek vengeance and to evoke fear and distress. In *Kennedy*, the punishment rested on an arbitrary decision and failed to comport with the evolving standards of decency.

Proportionality-based dignity violations mark an important but limited role in Eighth Amendment caselaw with respect to dignity claims. As I have shown, *Trop* and *Kennedy* articulated dignity as a constraint on punishments that are categorically excessive relative to the crime. Such invocations, however, remain comparatively rare in the Court's history, especially when contrasted with culpability-based and treatment-based dignity violations. It is in these latter contexts that dignity becomes a recurring justification in standards of punishment, both in terms of sentencing and in terms of prison conditions, and it is to these contexts that I now turn.

[574] *Kennedy v. Louisiana*, 554 U.S. 407, 439 (2008)

[575] *Kennedy v. Louisiana*, 554 U.S. 407, 439 (2008)

4. Type II: Culpability Dignity Violations ca. 1971-2004

In 1967, William Furman was burglarizing a home when he encountered one of its residents. In his attempt to flee, he tripped and fell. His gun discharged and killed the resident. In 1968, he was convicted of murder and sentenced to death after a one-day trial.[576] His subsequent appeal was decided by the Supreme Court alongside *Jackson v. Georgia* and *Branch v. Texas* in the landmark ruling, *Furman v. Georgia*. The question put before the Court was whether the death penalty violated the Eighth and Fourteenth Amendments.

Prior to *Furman*, juries in first-degree murder trials held virtually unbridled discretion when recommending the death penalty for a defendant. In a long and expansive ruling, the Court focused on the arbitrary and discriminatory imposition of the death penalty, including its disproportionate application to black defendants. Its decision effectively placed a moratorium on the death penalty while states designed more objective standards and procedures that would limit the arbitrariness of the death penalty.

One of *Furman's* striking features is that dignity plays a central role in its reasoning, a role most evident in Justice William Brennan's concurring opinion. Brennan began his opinion by invoking *Trop v. Dulles,* which he interpreted as grounding punishment in "the basic prohibition against inhuman treatment," a prohibition underwritten by "nothing less than the dignity of man."[577] On Brennan's reading, the Cruel and Unusual Punishment clause does not merely prohibit "uncivilized and inhuman punishments," but requires that the state, even when punishing, must treat "its members with respect for their intrinsic worth as human beings. A punishment is "cruel and unusual," therefore, if it does not comport with human dignity."[578] Although Brennan acknowledged that his formulation did not itself determine the constitutionality of specific punishments, he nevertheless articulated a set of principles "sufficient to permit a judicial determination whether a challenged punishment comports with human dignity."[579] The primary principle is that punishment must not be so severe as to degrade human dignity. Although he noted that considerations of pain are relevant, Brennan echoed *Trop*

[576] See Joan M. Cheever, *Dead Man Walking* (Wiley Press, 2006).

[577] *Furman v. Georgia*, 408 U.S. 238, 270 (1972)

[578] *Furman v. Georgia*, 408 U.S. 238, 270 (1972)

[579] *Furman v. Georgia*, 408 U.S. 238, 270 (1972)

by emphasizing that degradation extends beyond physical suffering.[580] A punishment that violates dignity does more than inflict pain; it treats certain humans as nonhuman, "as objects to be toyed with and discarded."[581] Even "the vilest criminal" Brennan insisted, remains a human being "possessed of common human dignity," and thus warrants standards of punishment.[582]

The remaining three principles flow from the first. A punishment fails to comport with dignity when it is arbitrarily imposed, when it is unacceptable to contemporary society, and when it is so excessive as to serve no purpose beyond the pointless infliction of suffering.[583] Brennan offered torture as a paradigmatic example, though he acknowledged that punishments rarely violate just one principle in such a blatant manner. In other words, punishments in the United States are seldom so obviously severe, arbitrary, socially rejected, or unnecessary as to violate dignity on the ground of one principle alone.[584] In practice, Brennan observed, the Court's cases tended to implicate several of these principles simultaneously and to a lesser degree. For instance, although it is clear that many Americans oppose the death penalty, it is less clear whether enough oppose it such that it would *de facto* violate the third principle; however, when we consider wavering social consensus alongside the arbitrary infliction of the death penalty, due in part to the punishment's relative rarity, the claim that the death penalty violates dignity becomes legally intelligible. It is therefore the application of the principles *in combination* that "provide means by which a court can determine whether a challenged punishment comports with human dignity."[585]

Furman v. Georgia did not end the death penalty, though it did rule that its arbitrary and inconsistent imposition violated the Eighth and Fourteenth Amendments. As Justice Stewart wrote, "These death sentences are cruel and unusual in the same way that being struck by lightning is cruel and unusual."[586] The Constitution could not "permit this unique penalty to be

---

[580] *Furman v. Georgia*, 408 U.S. 238, 271 (1972)

[581] *Furman v. Georgia*, 408 U.S. 238, 272-73 (1972)

[582] *Furman v. Georgia*, 408 U.S. 238, 273 (1972)

[583] *Furman v. Georgia*, 408 U.S. 238, 271-79 (1972)

[584] *Furman v. Georgia*, 408 U.S. 238, 281 (1972)

[585] *Furman v. Georgia*, 408 U.S. 238, 281 (1972)

[586] *Furman v. Georgia*, 408 U.S. 238, 309 (1972)

so wantonly and so freakishly imposed."[587] As a result, 629 death sentences were vacated and reduced to life imprisonment of which 322 were paroled.[588] However, *Furman's* status remains ambiguous. The landmark ruling brought renewed attention to a punitive practice that had been slowly withering away. With renewed attention came backlash from an American constituency that fervently reaffirmed its support for capital punishment.[589]

Following *Furman*, thirty-five states rushed to redraft their laws in an attempt to comply with the ruling all while preserving the death penalty.[590] Some states proposed solutions to nonarbitrary death sentences by imposing mandatory death sentences for all first-degree murders, as in the case of *Woodson v. North Carolina* in 1976. The Supreme Court would eventually strike down North Carolina's law by reasoning that it departed from social consensus against mandatory death sentences and also failed to provide juries with standards. Nevertheless, the Court affirmed the constitutionality of the death penalty so long as juries were able to consider the aggravating factors of the crime against the mitigating factors of the offender. The Court reaffirmed the death penalty that same year in a group of consolidated cases known as *Gregg v. Georgia*, in which it held that the death penalty was unconstitutional only when it failed to provide objective standards for juries in their deliberations.

Whereas many of the justices ruled in favor of the death penalty in *Furman*, albeit with considerable restrictions on its application, only Brennan drew on dignity to guide his decision. There, he notably opposed the death penalty under all and any circumstances. By 1976, however, the Court drew from Brennan's reasoning when considering capital punishment's comportment with the "basic concept of human dignity." [591] This time it affirmed the death penalty respected human dignity. On this point, the Court held capital punishment to be an appropriate expression of the community's belief that "certain crimes are themselves so grievous an affront to humanity

---

[587] *Furman v. Georgia*, 408 U.S. 238, 310 (1972)

[588] *"Furman v. Georgia*," NAACP Legal Defense and Educational Fund, Effective January, 2026, https://www.naacpldf.org/case-issue/furman-v-georgia/#/.

[589] For an overview of *Furman's* complicated effects, Carole Steiker and Jordan Steiker, *Courting Death: The Supreme Court and Capital Punishment* (Harvard University Press, 2016).

[590] Elaine McArdle, "The End of the Death Penalty: 'Unintended consequences' and the legacy of *Furman v. Georgia*" in *Harvard Law Bulletin* (2023).

[591] *Gregg v. Georgia*, 428 U.S. 153, 182-83 (1973)

that the only adequate response may be the penalty of death."[592] Everyone possesses human dignity, though some crimes shock the conscience in a way that warrants extreme punishments. In other words, some are more and less worthy of their dignity depending on their actions. So long as the administration of the punishment was procedurally fair and that the state only executed people who had committed the most shocking crimes, the Court saw no violation of dignity.

In his dissent, Brennan offered a more politically and morally robust sense of dignity than the one found in the Court's opinion. This suggests that, in 1976, there was a struggle over what it meant to respect a prisoner's dignity. In *Gregg,* Brennan referred to dignity as one of the "moral concepts" inherent in the Clause, a concept intended to ensure the state treat its citizens "in a manner consistent with their intrinsic worth."[593] Even when considering the vilest crimes, Brennan stressed "a punishment must not be so severe as to be degrading to human dignity. A judicial determination whether the punishment of death comports with human dignity is therefore not only permitted but compelled by the Clause."[594] Five years after *Furman*, most of the Court agreed with Brennan's reasoning and drew on his four principles in reaching its decision. The disagreement therefore concerned the threshold that dignity set within the evolving standards of decency and the appropriate practices associated with respecting an offender's dignity.

This is why *Woodsen v. North Carolina* specifies that respect for human dignity "requires consideration of aspects of the particular offense as a constitutionally indispensable part of the process of imposing the ultimate punishment of death."[595] The problem with North Carolina's mandatory death penalty, then, was that it treated the offenders as a "faceless, undifferentiated mass" and subjected the offender to uniform standards regardless of the unique features of his crime and of his individual circumstances.[596] Mechanistically handing down a death sentence without first attending to the particularity of the crime and the offender, and without allowing for

[592] *Gregg v. Georgia*, 428 U.S. 153, 184 (1973)

[593] *Gregg v. Georgia*, 428 U.S. 153, 229 (1973)

[594] *Gregg v. Georgia*, 428 U.S. 153, 229-30 (1973)

[595] *Woodsen v. North Carolina*, 428 U.S. 280, 281 (1976)

[596] *Woodsen v. North Carolina*, 428 U.S. 280, 281 (1976)

"the possibility of compassionate or mitigating factors stemming from the diverse frailties of humankind" failed to treat the prisoner with the dignity he deserved.[597]

The Court's emphasis on culpability is further illustrated in *Roper v. Simmons*, where it considered whether the execution of minors violated the Cruel and Unusual Punishment clause. In 1993, Christopher Simmons was sentenced to death for a crime committed when he was sixteen. During the mitigation phase of his trial, the Court noted that testimony was offered by corrections officers in the Missouri juvenile justice system as well as by his mother, father, and two younger half-brothers, a neighbor, and a friend, all of whom spoke of the close and loving relationships they had with Simmons.[598] Simmons' defense counsel also spoke of his young age and argued such a factor should make "a huge difference" to jurors when recommending the appropriate punishment.[599] Despite these mitigating circumstances, the jury sentenced Simmons to death.

In its ruling, the Court held that the execution of juveniles violated the Eighth Amendment. and categorically exempted minors from capital punishment.[600] The Court reasoned that juveniles are comparatively immature and irresponsible, making their conduct less morally reprehensible than that of adults.[601] Juveniles are not fully formed characters and so are more vulnerable to outside influences. They lack the maturity that would render them fully responsible for their crimes.[602] For these reasons, the Court held juveniles were less blameworthy for their actions, a ruling which extended the reasoning developed in *Atkins v. Virginia*, which prohibited the execution of intellectually disabled offenders. In both cases, the mitigating factors—immaturity, irresponsibility, and diminished rational capacity—are so substantial as to undermine full moral culpability. Such mitigating considerations *de facto* outweighed any aggravating circumstances. To disregard this would fail to respect the dignity of the offender.

---

[597] *Woodsen v. North Carolina*, 428 U.S. 280, 281 (1976)

[598] *Roper v. Simmons*, 543 U.S. 551, 558 (2005)

[599] *Roper v. Simmons*, 543 U.S. 551, 558 (2005)

[600] *Roper v. Simmons*, 543 U.S. 551, 598 (2005)

[601] *Roper v. Simmons*, 543 U.S. 551, 651 (2005)

[602] *Roper v. Simmons*, 543 U.S. 551, 561 (2005)

In all of the cases above, when the Court considered whether a punishment comported (or failed to comport) with human dignity, it considered the proportionality of the punishment relative to the offender. Although the Court held and reaffirmed that the death penalty does not violate human dignity, it nonetheless specified that, depending on the facts of the offense and the character of the offender, some individuals are more responsible and blameworthy than others. To respect someone's dignity, on this view, is to hold him accountable. It is to treat him as someone who ought to take responsibility for his actions, provided he has the capacity to do so. In the United States during the 1970s, social consensus appeared to support the view that the "worst" offenders warranted capital punishment. Yet, determining who is in fact blameworthy enough to warrant such punishment is not an easy task nor should it be an arbitrary decision. Respecting the dignity of someone convicted of first-degree murder, someone who has offended society's standards of decency, demands fair procedures and individualized consideration. Capital punishment should not be mechanistically imposed nor meted out without compassion for the offender. But neither should it be administered without a careful assessment of culpability grounded in clear and objective standards.

5. Type III: Treatment Dignity Violations ca. 1967-1974

In 1968, the U.S. Court of Appeals for the Eighth Circuit banned the use of corporeal punishment (otherwise known as "the strap,") as a disciplinary measure in Arkansas state prisons. At the time injunctive relief was sought, the court found that there were no written rules governing whipping and that punishment was subject only to an informal restraint of ten blows for a single offense.[603] In the lower courts, and prior to *Jackson v. Bishop's* landmark ruling, Judge Henly was hesitant to rule that corporeal punishment was per se unconstitutional, though he called for "appropriate safeguards" and "recognizable standards," for any future use of corporeal punishment. In response, prison officials implemented minimal standards, such as making the punishable offenses public and prohibiting corporeal punishment on bare skin. Even with these minimal standards in place, the Eighth Circuit was skeptical that any standards would successfully prevent the abuse of corporeal punishment. In the hands of "the sadistic and the unscrupulous"

[603] *Jackson v. Bishop*, 404 F.2d 571, 574 (1968).

these standards would go unobserved.[604] More to the point, the court noted the difficulty in determining the precise moment at which a permissible form of whipping turned impermissible.

In its reasoning against corporeal punishment, the court followed *Trop v. Dulles's* by affirming that the basic concept underlying the Eighth Amendment "is nothing less than the dignity of man."[605] It also provided a more substantive reading of dignity, one that would later inform Supreme Court reasoning. In *Jackson*, the court emphasized the centrality of "man's basic dignity," to the Eighth Amendment, noting that it embodies the "broad and idealistic concepts of dignity, civilized standards, humanity, and decency."[606] The ruling also observed that the strap undermines prisoners' sense of self-respect, generates resentment that may extend toward the officer administering the punishment and towards the prison institution itself. Moreover, it asserted that no justification could support corporeal punishment and wrote that "human considerations and constitutional requirements are not, in this day, to be measured or limited by dollar considerations or by the thickness of the prisoners' clothing."[607]

Once *Jackson v. Bishop's* ruling established precedent for federal courts to rule on the constitutionality of treatment within prisons (and not just in sentencing), four requests for injunctive relief arrived in quick succession: *Holt v. Sarver I* (1969) and *Holt v. Sarver II* (1970), *Landman v. Royster* (1971), and *Gates v. Collier* (1972). The prisons under scrutiny shared traits that were common to prisons at the time. They effectively operated as plantations and they were unhygienic, improperly heated, employed few prison staff, and instead relied on armed prisoner "trusties" to perform custodial duties over other prisoners.[608]

The claims of both *Holt* I and II concerned basic living conditions. Prisoners slept in open barracks, leaving them insecure and susceptible to attack. When placed in punitive segregation, they experienced unhygienic conditions. The court determined that its oversight did not end with sentencing. Cruel and unusual punishment could be found in the very conditions of confinement,

---

[604] *Jackson v. Bishop*, 404 F.2d 571, 579 (1968)

[605] *Trop v. Dulles*, 356 U.S. 86, 100 (1958)

[606] *Jackson v. Bishop*, 404 F.2d 571, 574 (1968)

[607] *Jackson v. Bishop*, 404 F.2d 571, 579 (1968)

[608] *Holt v. Sarver*, 309 F. Supp. 362, (E.D. Ark. 1970)

conditions "so bad as to be shocking to the conscience."[609] The most egregious type of cruelty, the court stated, was the trustie system. Under such conditions, "a convict, however cooperative and inoffensive he may be, has no assurance whatever that he will not be killed, seriously injured, or sexually abused… Apart from the physical danger, confinement in the Penitentiary involves living under degrading and disgusting conditions."[610] The court continued, "For the ordinary convict a sentence to the Arkansas Penitentiary today amounts to a banishment from civilized society to a dark and evil world completely alien to the free world, a world that is administered by criminals under unwritten rules and customs completely foreign to free world culture."[611] Following its denouncement of the conditions at Arkansas state prison, the court ordered a "prompt and reasonable start toward eliminating the conditions," writing, "The lives, safety, and health of human beings, to say nothing of their dignity, are at stake. The start must be prompt, and the prosecution must be vigorous."[612]

One year later, the Virginia Penal system was under the scrutiny of the courts in a class-action lawsuit in *Landman v. Royster.* After the prisoners initiated a peaceful work stoppage in 1968, prison officials responded with retaliatory discipline such as prolonged starvation through a bread and water diet with calories amounting to little more than 700 per day. In its ruling, the court noted that although there were few standards to distinguish between permissible penalties and ones that were cruel and unusual, it drew its reasoning from *Trop,* in which "the basic concept underlying the Eighth Amendment is the dignity of man."[613] From this remark, the court elaborated that a bread and water diet does not only physically debilitate a prisoner, amounting to "a prolonged sort of corporeal punishment," but weakens the prisoner's mental alertness.[614] In addition, the ruling noted that prolonged starvation only served to remind the prisoner that the prison authorities "enjoy complete control over all sources of pleasure, comfort, and basic

[609] *Holt v. Sarver*, 309 F. Supp. 362, 364 (E.D. Ark. 1970)

[610] *Holt v. Sarver*, 309 F. Supp. 362, 381 (E.D. Ark. 1970)

[611] *Holt v. Sarver*, 309 F. Supp. 362, 381 (E.D. Ark. 1970)

[612] *Holt v. Sarver*, 309 F. Supp. 362, 371 (E.D. Ark. 1970)

[613] *Landman v. Royster,* 333 F. Supp. 621, 646 (E.D. Va. 1971); *Trop v. Dulles*, 356 U.S. 86, 100-101 (1958)

[614] *Landman v. Royster,* 333 F. Supp. 621, 647 (E.D. Va. 1971)

needs."[615] It thus concluded that, "the bread and water diet is inconsistent with current minimum standards for respect of human dignity" and as such, violated Eighth Amendment.

In a third landmark ruling from the lower courts, *Gates v. Collier*, prisoners at Mississippi State Penitentiary, also known as Parchman Farm contested the conditions of their confinement. At the time of the lawsuit, Parchman farm contained 1900 prisoners, 50 of which were women, and two-thirds were black.[616] Ruling in their favor, the court noted that "[t]he defendants have subjected the inmate population at Parchman to cruel and unusual punishment by failing to provide adequate protection against physical assaults, abuses, indignities and cruelties of other inmates," through a totality of conditions.[617] For example, although it might not be cruel and unusual to place a prisoner in solitary confinement, it is so if this occurs in conjunction with basic elements of hygiene, if the prisoner is placed on a starvation diet for indeterminate periods of time, or if he has been subjected to other brutal and degrading practices in the course of solitary confinement. As with *Landman*, the ruling in *Gates* takes a more expansive approach to corporeal punishment than simply a prohibition on whipping. In one of the clearest, and most specific rulings on dignity vis-a-vis the Cruel and Unusual Punishment clause, the court declared that the defendants were:

> prohibited from imposing any form of corporeal punishment of such severity as to offend present-day concepts of decency and human dignity. Specifically, defendants are enjoined from punishing inmates by methods previously used at the penitentiary and found by this court to be excessive: beating; shooting; administering milk of magnesia; stripping inmates of their clothes; turning fans on inmates while they are naked and wet; depriving inmates of mattresses, hygienic materials and/or adequate food; handcuffing or otherwise binding inmates to fences, bars, or other fixtures; using a cattle prod to keep inmates standing or moving at Maximum Security Unit (MSU) or elsewhere; shotting at or around inmates to keep them standing or moving while at MSU; forcing inmates to stand, sit or lie on crates, stumps or otherwise maintain awkward positions for prolonged periods.[618]

In 1976, after seven years of litigation and court injunctions, the Arkansas state prison system was under scrutiny yet again, this time at the level of the Supreme Court with *Hutto v.*

---

[615] *Landman v. Royster,* 333 F. Supp. 621, 647 (E.D. Va. 1971)

[616] *Gates v. Collier*, 349 F. Supp. 881, 886 (N.D. Miss. 1972)

[617] *Gates v. Collier*, 349 F. Supp. 881, 894 (N.D. Miss. 1972)

[618] *Gates v. Collier*, 349 F. Supp. 881, 899-900 (N.D. Miss. 1972)

Finney. In 1970, the Arkansas Supreme Court had found Cummins Farm to be overcrowded. By 1976, the population had increased to 1,500. Practically all prisoners were losing weight on the prison's "grue" diet (i.e., mashed meat, potatoes, oleo, syrup, vegetables, eggs, and seasoning baked in a pan).[619] Although significant improvements had been made in the years following Holt *I* and *Holt II*, conditions again deteriorated once the District Court had withdrawn its supervision in 1973.

In its preliminary remarks, the Supreme Court noted that the punitive isolation cells were "particularly disturbing," with twice as many prisoners as beds in some cells and an inadequate number of guards, who frequently assaulted the prisoners with nightsticks and mace.[620] The Court held that the Eighth Amendment's ban on inflicting cruel and unusual punishment extended out from "physically barbarous punishments" to ones that are "grossly disproportionate to the offense," and that transgress the "broad and idealistic concepts of dignity, civilized standards, humanity and decency."[621] When determining whether punitive isolation of over thirty days violated human dignity and thus warranted protection by the Eighth Amendment, the Court noted that the constitutionality of punitive isolation should not be considered in a "vacuum."[622] In other words, though punitive isolation is in itself permissible, the punishment violates the Eighth Amendment when that isolation is set against a totality of conditions of cruelty. When the Court considered punitive isolation *alongside* poor diet, continued overcrowding, violence and beatings, vandalized cells, and the lack of professional standards on the part of the prison guards, it was clear that the totality of conditions amounted to cruel and unusual punishment.[623] As the Court stated, "A filthy, overcrowded cell and a diet of 'grue' might be tolerable for a few days and intolerably cruel for weeks or months."[624] In this case, it is not simply the totality of conditions that amount to cruel and unusual punishment, but how the cruelty of these conditions amount to physical and psychological torture over a prolonged period.

---

[619] *Hutto v. Finney*, 437 U.S. 678, 683; 678 (1978)

[620] *Hutto v. Finney*, 437 U.S. 678, 684 (1978)

[621] *Hutto v. Finney*, 437 U.S. 678, 685 (1978)

[622] *Hutto v. Finney*, 437 U.S. 678, 678 (1978)

[623] *Hutto v. Finney*, 437 U.S. 678, 678 (1978)

[624] *Hutto v. Finney*, 437 U.S. 678, 677-78 (1978)

What emerges from these cases is a substantive set of practices and conditions that violate the dignity of prisoners.[625] As in the case of capital punishment, it is rarely a single condition that constitutes a violation. Rather, the totality of conditions creates an environment that inflicts physical and psychological harm, degrades a prisoner's sense of self-respect and safety, and erodes trust. Under these conditions, the prisoner has no assurance that he will escape the prison unscathed. Central to these violations is the evident absence of standards, which encourages cruelty and undermines human dignity. At least in the contexts of prisons, respect for dignity does not entail access to a private cell, proper nutrition, or a prohibition on solitary confinement. It does, however, set a threshold against conditions of overcrowding, starvation, and indefinite punitive segregation. In other words, dignity provides a minimal standard within institutions that are otherwise prone to abuse through unbridled discretion.

6. The Emerging Rights of the Confined

For most of the twentieth century, federal courts adhered to what is now known as the "hands-off" doctrine of prison administration, under which federal oversight of correctional practices rarely occurred. Prisoners were, in effect, "slaves of the state." [626] From our present vantage point, it is easy to forget that the Cruel and Unusual Punishment clause did not apply to state prisons until 1962 with *Jones v. Cunningham*. After *Jones*, prisoners succeeded in challenging the conditions of their confinement, as my discussion above has demonstrated.

Judicial rulings are not inert documents. They are legally binding and normatively efficacious constraints on our practices. Accordingly, a genealogy must understand how correctional institutions articulated these emerging Supreme Court rulings as well as how the standards the courts set down as legally binding were translated into training techniques for prison officials. Conversely, a genealogy must understand how these standards created by prison administrators informed and influenced court rulings. To do so, I focus on three training manuals published between 1966 and 1973. The first is the American Correctional Association's *Manual for Correctional Standards* (1966), alongside its corresponding *Study Guide for the Application of the Manual of Correctional Standards* (1970), a guide intended as a tool to evaluate and

[625] For a related reading of the totality of conditions violating human dignity, see Eric G. Woodbury, "Prison Overcrowding and *Rhodes v. Chapman*: Double-Celling by What Standard?" *Boston Law Review* 23: 713-760.

[626] *Ruffin v. Commonwealth*, 21 Gratt. 790, 796 (1871)

ensure appropriate compliance with the principles and procedures set out in the manual. The second is *The Emerging Rights of the Confined* (1972), written by William D. Leeke, then Commissioner of the South Carolina Department of Corrections. The third is *The Jail: Its Operations and Management* (1973), published by the United States Bureau of Prisons, which set forth standardized procedures and principles for jail administration. These manuals, I suggest, exemplify both a body of knowledge of standards of "dignified treatment," as well as a set of procedures and techniques that function as expressions of those standards.

In *Discipline and Punish*, Foucault characterizes disciplinary power as a series of procedures and strategies that train the body through carefully defined habits and routines. Disciplinary power produces analyzable and manipulable bodies, renders them docile, and operates by targeting the body's most mundane elements through continuous and uninterrupted exercise. Foucault primarily focuses on how disciplinary power transforms the delinquent (alongside other members of society) into a docile body and pays comparatively little attention to how correctional officers themselves are also subjected to disciplinary mechanisms. In examining these manuals, my aim is therefore to consider how disciplinary power functions to train and produce the "correct" conduct of prison officials. I analyze this process across three strategic levels. First, the psychological dispositions that orient officers toward dignified treatment of prisoners, such as institutional paranoia and rehabilitative concern. Second, the techniques that act directly on officers' bodies through prescribed routines and the gradation of tasks. Third, the instruments through which standards of treatment are formalized and evaluated, such as the study guide and accreditation form.

Psychological Dispositions. In 1972, William D. Leeke, then commissioner of the South Carolina Department of Corrections, catalogued and interpreted over 1,000 State Supreme Court, Federal District Court, and United States Supreme Court decisions for the purpose of overcoming the "hostility born of paranoia" that had overtaken state prisons in the wake of the growing prisoner's rights movement.[627] Alongside the landmark rulings in favor or prisoner rights, Attica prison had witnessed the deadliest prison riot in United States history. Though Attica is notably left out of the Leeke's *Emerging Rights of the Confined*, one can reasonably

[627] Wiliam D. Leeke, *The Emerging Rights of the Confined* (The Correctional Development Foundation, 1972), 26.

infer the wish to prevent greater violence, alongside the fear of administrators being named as defendant in the suit, fed into the atmosphere of paranoia.[628]

In the chapter "Disciplinary Methods," Leeke draws from Justice Warren's opinion in *Trop v. Dulles* to guide his interpretation of the Eighth Amendment Cruel and Unusual Punishment clause. In that ruling Justice Warren noted that even though the basic concept underlying the Eighth Amendment is the dignity of man, the Cruel and Unusual Punishment clause derives its precise meaning from evolving standards of decency.[629] Despite the evolving meaning of the Cruel and Unusual Punishment clause, Leeke nevertheless recognized "the need for definite rules," in order to curb the "unfettered discretion" that had defined the prison system for much of the twentieth century.[630] No longer, on this view, could prison administrators arbitrarily or capriciously inflict punishment on those they confined. What is more, Leeke argued that the burden had shifted to prison administrators to justify restrictive policies by showing that they served a legitimate state interest and imposed the minimum restriction necessary to achieve their purpose. "No longer," Leeke wrote, "must the inmate establish that he is entitled to the right; that is generally assumed today. The burden is now squarely on the administrator to establish that the particular restriction on the exercise of the right is absolutely necessary to the maintenance of the compelling interest of prison security. And it must be remembered that this burden can no longer be made by mere opinion or speculation."[631] To illustrate his analysis, Leeke pointed to a case in which prison officials failed to meet this burden after denying a prisoner work privileges on the basis of their "opinion" that the prisoner was "unsuitable" for work.[632] These words, according to Leeke, operate as "catch-all phrases," that no longer served as an appropriate justification. Other phrases he identified as insufficient justifications included appeals to "the good of the institution," "inappropriate conduct," and "negative attitude," insofar as such phrases functioned as thinly veiled attempts to legitimate punishment on the whim of the official.[633] If a prison

[628] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 26.

[629] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 102.

[630] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 104.

[631] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 30.

[632] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 34

[633] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 34-35.

administrator wished to avoid litigation, or successfully defend against it, he needed to provide a definite and articulable basis for punishment.

There were reasons beyond the fear of litigation that led prison administrators to call for definite rules, chief among them was the related concern about security and control. Indeed, even prior to some of the landmark prisoners' rights rulings, the American Correctional Association had already called for minimum standards of correctional administration. As the association noted in its manual, "both legislators and courts have been reluctant to limit the discretion of administrative authorities… As a practical matter, therefore, convicted offenders—and prisoners in particular—are largely at the mercy of correctional authorities."[634] In response, the American Correctional Association set out "policy standards to guide practice in the absence of legislation," which they explicitly drew from the Standard Minimum Rules for the Treatment of Prisoners adopted by the United Nations.[635] Notably, the first "basic principle," of the United Nations's rules maintain that "all prisoners shall be treated with the respect due to their inherent dignity and value as human beings" and that no prisoner shall be subjected to torture or other cruel, inhumane or degrading treatment under any circumstance.[636]

What this indicates is that although many of the manuals published during this period spoke to the anxieties generated by federal court intervention, the adoption of dignity as an animating principle was not driven by judicial pressure alone. Indeed, the courts often cited standards and manuals when making their rulings.[637] The American Correctional Association's call for standards was therefore grounded in pragmatic concerns about institutional security. Prisoners who are adequately fed and clothed, provided with meaningful work, and given avenues for self-improvement are less likely to attempt escape than those who are subjected to prolonged inactivity, moral degradation, and humiliation. Increased security also created the conditions for rehabilitation, which was yet another reason why prison officials called for increased standards. In *The Jail: Its Operation and Management*, for example, the manual expressed a need for jail administrators to change their attitudes from unbridled entitlement to a concern for rehabilitation.

---

[634] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1966), 129.

[635] American Prison Association, *A Manual of Correctional Standards* (American Prison Association, 1966), 129.

[636] UN General Assembly Resolution 70/175 (2015), *United Nations Standard Minimum Rules for the Treatment of Prisoners (the Nelson Mandela Rules*).

[637] Sterling v. Cupp, 625 P.2d 123 (Or. 1981).

Such an attitude constituted "an essential ingredient of any rehabilitative program is the preservation of individual dignity."[638] Part of what it means to engage in a dignified rehabilitative program is to ensure that the prisoner was given reasons for his punishment and meaningful avenues to contest that punishment.[639]

Techniques. Having the appropriate dispositions to endorse standards of justified punishment is not, on its own, sufficient to explain how dignified treatment became operationalized within prison corrections. For this, we require further elaboration of what Foucault calls *techniques* or a "micro-physics of power."[640] One such technique is that of the exercise, which trains the body through the gradation of tasks. A paradigmatic exercise in prison administration is frisking, to which prisoners are routinely subjected and which was standardized in correctional manuals. These standards expressed an explicit aim of preserving a prisoner's dignity during the procedure. For example, a jail training manual published by the National Institute of Corrections in 1986 noted that "most persons are concerned about maintaining their personal dignity. The administration process can be embarrassing and difficult. When possible, jail personnel should do whatever they can do to permit the inmate to maintain his privacy and dignity when he is bathing and during the search."[641] These general directives were often accompanied by detailed instructions that specify a step-by-step procedure and illustration. There should be no other prisoners around. The prisoner should empty his pockets. The administrator should have the prisoner turn to the wall and spread his legs with his arms extended. Once the position is assumed, the administrator should scan the prisoner using both his hands, beginning at the hairline, then proceeding down the back of the head, neck, chest, stomach, back, beltline, legs, crotch, and so forth. Cross-gender frisks and strip searches should be avoided at all cost and internal rectal or vaginal checks should only be carried out if there is reasonable cause that the prisoner is carrying contraband.[642]

---

[638] Nick Pappas, ed., *The Jail: Its Operation and Management* (United States Bureau of Prisons, 1973), 187.

[639] Wiliam D. Leeke, *The Emerging Rights of the Confined*, 131.

[640] Michel Foucault, *Discipline and Punish* (Vintage Books, 1975), 26.

[641] National Sherrif's Association*, Jail Officer's Training Manual* (National Institute of Corrections, U.S. Department of Justice, 1986), 171.

[642] National Sherrif's Association*, Jail Officer's Training Manual* (National Institute of Corrections, U.S. Department of Justice, 1986), 167.

The manuals frequently describe prison administrators as acting in service of the state's power: "the administrator *personally* is not the real defendant in these suits. It is the exercise of state power that is being challenged, and the administrator is named in the suit as defendant merely because he is the instrument of the state power that is asserted."[643] If the prison administrator is represented as an embodied instance of state power, then frisking emerges as one of the most intimate points of contact between the state and the prisoner. The state already subjects the prisoner to complete control over his quotidian habits of eat, sleep, work, and exercise. If this control were not yet apparent, frisking serves to remind the prisoner that there is no part of the body that the state does not touch. And yet, standards governing frisking can curtail the unbridled discretion that might otherwise occur in the name of security or control. Procedures that are improperly followed, or that exceed their prescribed limits, intensify the humiliation that is already constitutive of frisking. The routinization of frisking thus creates a distance between the officer and the act itself. It also establishes ways of touching that reach beyond the pale of justification, for example, if a male officer performs a vaginal examination. In such a case, the prisoner will not only have suffered an indignity, but she will be entitled to requesting injunctive relief.

To be clear, the act of frisking was not contested by the courts or the public, as it was (and still is) widely accepted that a minimum degree of humiliation—or "embarrassment," as the manuals often describe it—is to be expected in service of institutional safety and security. Nevertheless, administrators were prohibited from frisking in a way that invaded "the prisoner's residuum of personal dignity" insofar as the treatment "goes beyond recognized necessity."[644] The procedures governing frisking thus exemplify dignity's point: to draw a boundary between what is penologically necessary and therefore justified and what is undue and unnecessary. Although punishment might be harsh, degrading, and dehumanizing insofar it entails the loss of liberty and basic comforts, dignity functions to limit forms of harsh, degrading, and dehumanizing treatment that are gratuitously so. In the case of frisking, the prisoner, though examined, is shielded from the gaze of other prisoners. Sexual exploitation is mitigated through the prohibition of cross-gender frisking. The exercise becomes a routinized check, in which the

[643] Wiliam D. Leeke*, The Emerging Rights of the Confined*, 32-33; Pappas, *The Jail: Its Operation and Management*; *Draft Federal Standards for Corrections*, 45.

[644] Sterling v. Cupp, 625 P.2d 123 (Or. 1981).

prisoner assumes a prescribed stance and the official performs as standardized and automatic sweep of the body as a machine might.

Instruments. The manuals called for definite rules and regulations upon which punishment is based. But as Leeke observes in *The Emerging Rights of the Confined*, it was not enough to establish rules and standards for prison administrators to follow: "The rules of the institution not only must exist in a definite form prior to the imposition of punishment, but they must be observed by the penal officials."[645] He goes on to quote *Landman v. Peyton*, in which the court stated "While written regulations governing [inmate] conduct have been promulgated, these are insufficient if no attempt is made to insure their observance."[646] The court thus placed an affirmative duty on prison administrators to ensure that standards were properly operationalized by lower-level officials, so as to prevent the arbitrary and capricious imposition of rules. One way in which these standards were operationalized was through the cultivation of dispositions, such as paranoia of litigation and security or rehabilitative concern. Another was through techniques, which prescribed the appropriate exercises for dignified treatment. There is yet a third way for us to consider, that of the evaluation, which exerts a normative force on all levels of prison officials through instruments of hierarchical observation and formalization.

In *Discipline and Punish*, Foucault observes that the success of disciplinary power depends on the instrument of the examination or evaluation, which is a highly ritualized procedure comprised of "a normalizing gaze, a surveillance that makes it possible to qualify, to classify and to punish."[647] Rather than the prisoner himself as the subject and object of evaluation, I suggest we can understand prison officials and the prison itself as another site for evaluation. Much like in Foucault's description, the evaluation of prison officials makes visibility and objectification compulsory, insofar as each area of corrections is prime for an observing gaze. Once administrators begin to enact the rules of corrections, they are captured and fixed into a network of documentation, expressed in notations, records, and files.[648] Through disciplinary writing, we

[645] Leeke, *The Emerging Rights of the Confined*, 105.

[646] *Landman v. Peyton* 370 F.2d 135 (4th Cir. 1966)

[647] Foucault, *Discipline and Punish*, 184.

[648] Foucault, *Discipline and Punish*, 189.

can begin to apprehend the formalization of prison conditions and prison treatment within power relations.

A clear example of the evaluation in prison corrections is the ACA's *Study Guide for the Application of the Manual of Correctional Standards*, which is a tool "for the evaluation of correctional programs, agencies and systems."[649] According to the authors, the evaluation reports in the study guide provided administrators with "a tool for self-evaluation" relative to the principles found in the *Manual of Correctional Standards*. They enabled administrators to graphically represent the relative strengths and weaknesses of their institution, and they include forms designed to involve all levels of staff as evaluators.[650] In the *Manual of Correctional Standards*, each chapter begins with a numbered listing of the basic principles, which is then followed by a discussion of the listed factors in paragraphs with corresponding numbers for the evaluation sheet. For the initial evaluation, a "supplemental rater," is expected to compare the principles found in the manual against the conditions of the prison through a simple rating system of "(x)" as an affirmative adherence to the standards, "(-)" as some minor various or omission of the standards, and "(0)" as the absence or ineffectiveness of standards. Once a supplemental rater has completed his evaluation, a "person responsible for final rating," scores the items on a chart which is then graphically represented.[651]

If we consider the evaluation form on "Legal Rights of Probationers, Prisoners, and Parolees," we can see that many of elements associated with dignified treatment, drawn from both state and federal case law as well as from the manuals themselves, were framed as standards against which evaluators were to measure their own institutional practices. Rules were required to be corrective rather than abusive or punitive; they must be minimal and no more restrictive than necessary to maintain orderly conduct; they must not be cruel inhumane, or degrading in nature; and corporeal punishment or instruments, such as clubs, were explicitly prohibited.[652] Through such evaluations, the prison administrator could compile a whole network of

---

[649] American Correctional Association, *Study Guide for the Application of the Manual of Correctional Standards* (American Correctional Association, 1970), 1.

[650] American Correctional Association, *Study Guide of the Application of the Manual of Correctional Standards*, 1.

[651] American Correctional Association, *Study Guide of the Application of the Manual of Correctional Standards*, 6.

[652] American Correctional Association, *Study Guide of the Application of the Manual of Correctional Standards*, Chapter 15.

documentation in which standards became formalized. With greater documentation, disciplinary power integrated each prison into a field of knowledge, so as to classify, rank, and determine its conformity to the legal norm.[653]

7. Putting Dignity First

In this chapter, I have presented a brief legal genealogy of dignity as a guiding principle within United States prisons. I have remained at the level of description, in the sense that I have aimed to explain dignity as a thick ethical concept in American jurisprudence and correctional practices. The point of this genealogy has been to give content to what is otherwise thought of as a formal, thin concept, and to show how, in practice, dignity has functioned as action-guiding. This means I have made sense of its history in relation to other ethical concepts, such as decency, safety, basic needs, humanity, responsibility, and culpability, alongside counter-ethical concepts, such as fear, cruelty, degradation, humiliation, and torture. I have also clarified dignity through the kinds of treatment it demands, including attention to the individual in the imposition of punishment, and through the kinds of treatment it prohibits, including forms of cruelty that few of us can readily imagine. With this descriptive account as a backdrop, I now conclude by briefly outlining a normative account of dignity, one that I offer as a hypothesis rather than a foregone conclusion.

In Chapter Two, I discussed Shklar's "cruelty principle," a principle that aims to mitigate the pervasive and unbridled fear that a state inflicts on its citizens. What is distinct about Shklar's normative approach is that it is oriented towards a *summum malum,* an idea which specifies that politics should be based on avoiding the worst possible outcome rather than pursuing a utopian *summum bonum*. In "The Liberalism of Fear," Shklar claims, "Liberals can begin with cruelty as the primary evil only if they go beyond their well-grounded assumption that almost all people fear it and would evade it if they could. If the prohibition of cruelty can be universalized and recognized as a necessary condition of the dignity of persons, then it can become a principle of political morality."[654] In the history of Eighth Amendment caselaw, dignity and cruelty appear in tandem, with each regulating punishment through their roles as a thick ethical concept and a

[653] See Foucault, *Discipline and Punish*, 190.

[654] Shklar, "The Liberalism of Fear," in *Liberalism and the Moral Life*, ed. Nancy L. Rosenblum (Harvard University Press: 1989), 21-38.

counter-ethical concept, respectively. Even though Shklar claims that mitigating cruelty is a necessary pre-condition for securing the dignity, her claim can also be read as articulating a morally austere, but politically ambitious normative framework for identifying the distinctive wrongs of cruelty. Cruelty is not wrong merely because it inflicts pain, suffering, fear, or distress, but because the treatment that produces such outcomes is incompatible with our conception of the minimum ethical standards owed to human beings. Likewise, within the genealogy I have offered, a dignity violation marks the point at which punishment falls below this minimum threshold of justified treatment and thereby gives rise to cruelty.

So then, what would it look like to formulate a principle of political morality that, by putting cruelty first, upholds the dignity of persons, and that, by affirming dignity, prohibits cruelty? I suggest that we construct a principle that constrains the two poles of legalism: on the one hand, there is an overly strict and mechanistic application of rules that leaves no room for compassion or attention to individual circumstances. On the other, there is a lack of clear standards and rules that would effectively constrain state power over the prisoner. I have shown that undefined guardrails on either side have led to some of the cruelest forms of treatment, treatments that are degrading to human dignity. Conversely, dignity has functioned as an effective concept for guarding against cruelty.

Consider, for example, that prior to *Furman v. Georgia*, juries were permitted to recommend the death penalty without clear or objective standards, a practice *Furman* deemed unconstitutional. In response, North Carolina imposed a mandatory death penalty for all first-degree murders, regardless of circumstances, a practice overturned in *Woodsen v. North Carolina*. In both cases, the Court ruled that dignity was violated, whether due to the absence of standards which resulted in arbitrary sentencing, or due to the rigid application of a rule that foreclosed compassion and individualized judgement. A similar dynamic appears in challenges to prison conditions, particularly in the 1970s, when the absence of clear standards left state prisons in abject conditions. Here again, dignity operated as a reason for instituting minimum standards of treatment, including adequate medical care and nutrition, prohibitions on physical abuse, and limits on the use of solitary confinement.

Dignity, of course, is an imperfect concept, much like other concepts, such as justice and equality. Yet, like justice and equality, it is not a concept we can, or should, dispense with in political or legal discourse. Instead, dignity can be understood as a strategy for action. Moving

forward we need to construct a normative theory that vindicates dignity in order to prohibit the worst forms of treatment a government can inflict on an individual. On this view, appeals to dignity would not present the concept as a permanent solution to the problem of cruelty, but as one that has pragmatically and strategically proven effective within specific legal and political contexts.

My point here is not to argue that those who offer a robustly moral account of dignity are mistaken. My claim is instead that dignity, understood in the thick sense I have developed in this chapter, does not reliably guide the kinds of action required to bring about those forms of treatment. Even Justice Brennan, one of the fiercest advocates of dignity in American jurisprudence, does not invoke the concept to justify claims that prisoners are entitled to green space or proper nutrition, as some more moralized accounts suggest. Instead, he argues that dignity constrains the state from executing its citizens, regardless of the offense committed. This marks a comparatively low threshold of treatment, one that as subsequent jurisprudence shows, even the Supreme Court or the wider public, has not fully endorsed. Of course, I believe that prisoners deserve much more than the prevention of cruelty, but that does not abrogate dignity's role in constraining the two poles of legalism. Alongside cruelty, our normative framework ought to put dignity first.

BIBLIOGRAPHY

Allen, Amy. *The Politics of Ourselves*. Columbia University Press, 2008.

Allen, Barry. "Another New Nietzsche." *History and Theory* 42, no. 3 (2003): 363-377.

American Correctional Association, *A Manual of Correctional Standards*. American Correctional. Association, 2018.

American Prison Association, *A Manual of Correctional Standards*. American Prison Association, 1954.

American Prison Association. *A Manual of Correctional Standards*. American Prison Association, 1966.

Anderson, Elizabeth. *Value in Ethics and Economics*. Harvard University Press, 1993.

———. "Human Interests, and Objectivity in Feminist Epistemology." *Philosophical Topics* 23, no. 2 (1995): 27-58.

———. "What Is the Point of Equality?" *Ethics* 109, no. 2 (1999): 287-337.

———. "Use of Value Judgements in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce." *Hypatia* 19, no.1 (2004): 1-24.

———. "Toward a Non-Ideal Relational Methodology for Political Philosophy: Comments on Schwartzman's *Challenging Liberalism*." *Hypatia* 2, no. 4 (2009): 130-145.

———. "The Fundamental Disagreement Between Luck Egalitarians and Relational Egalitarians." *Canadian Journal of Philosophy* 36 (2010): 1-23.

Appiah, Anthony. *The Honor Code: How Moral Revolutions Happen.* Norton Press, 2010.

Arendt, Hannah. *The Human Condition*. Chicago University Press, 1958.

———. "The Social Question." In *On Revolution.* Penguin Classics, 1963/2006.

Attica Liberation Faction. "Manifesto: The Attica Liberation Faction Manifesto of Demands and Anti-Depression Platform, 1971." *Black Agenda Report.* Effective April 13, 2026. https://blackagendareport.com/manifesto-attica-liberation-faction-manifesto-demands-and-anti-depression-platform-1971

Bartky, Sandra. *Femininity and Domination: Studies in the Phenomenology of Oppression*. Routledge Press, 1990.

Benhabib, Seyla. *Situating the Self: Gender, Community, and Postmodernism in Contemporary Ethics.* Taylor & Francis Group, 1992.

———. *Dignity in Adversity: Human Rights in Troubled Times*. Polity Press, 2011.

Bloch, Marc. *The Historian's Craft: Reflections on the Nature and Uses of History and the Techniques and Methods of Those Who Write It.* Translated by Peter Putnam. Vintage Books, 1949.

Bureau of Justice Statistics. *Prisoners in 1980.* U.S. Department of Justice, 1981: 2-3.

Cheever, Joan M. *Dead Man Walking*. Wiley Press, 2006.

Cherry, Myisha. *The Case of Rage: Why Anger Is Essential to Anti-Racist Struggle.* Oxford University Press, 2021.

Cohen, G.A. "Facts and Principles." *Philosophy & Public Affairs* 31, no. 3 (2003): 211-245.

Collingwood, R.G. *The Idea of History*. Oxford University Press, 1946.

Cottingham, John. "The Good Life and the Radical Contingency of the Ethical." In *Reading Bernard Williams*, edited by D. Callcut. Routledge Press, 2008.

Craig, Edward. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford University Press, 1990.

Crenshaw, Kimberlé. "Mapping the Margins: Intersectionality, Identity, Politics, and Violence Against Women of Color." *Standford Law Review* 43, no.6 (1991): 1241-1299.

Cudd, Ann. *Analyzing Oppression*. Oxford University Press, 2006.

———. "The Paradox of Liberal Feminism: Preference, Rationality, and Oppression," in *Varieties of Feminist Liberalism*, edited by Amy R. Baehr. Rowman & Littlefield, 2006.

Darwall, Stephen. *The Second Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press, 2009.

Davidson, Arnold. *The Emergence of Sexuality: Historical Epistemology and the Formation of Concepts.* Harvard University Press, 2001.

Davis, Angela and Elizabeth Martinínez, "Coalition Building Among People of Color." *Inscriptions* (7), 1994.

Debes, Remy. "Dignity's Gauntlet." *Philosophical Perspectives* 23, no. 1 (2009): 45-78.

Dillon, Robin S. "Respect and Care: Toward Moral Integration." *Canadian Journal of Philosophy* 22, no. 1 (1992): 105-131

Dreyfus, Hubert and Paul Rabinow. *Michel Foucault: Beyond Structuralism and Hermeneutics.* 2nd edition. The University of Chicago Press, 1982.

———. "What is Maturity?" In *Foucault: A Critical Reader,* edited by David Hoy. Wiley-Blackwell, 1991.

Erlenbusch-Anderson, Verena. *Genealogies of Terrorism*. Columbia University Press, 2018.

Estlund, David. *Utopophobia: On the Limits (If Any) of Political Philosophy*. Princeton University Press, 2019.

Etinson, Adam. "What's So Special About Human Dignity?" *Philosophy & Public Affairs* 46, no.4 (2020): 353-381.

Farrell, Colin. "Justice in Ideal Theory: A Refutation." *Political Studies* 55, no. 4 (2007): 844-864.

Faubion, James, ed., *Essential Works of Foucault Vol. 2*. The New Press, 1998.

Forrester, Katrina. "Judith Shklar, Bernard Williams and Political Realism." *European Journal of Political Theory* 11, no. 3 (2012): 247-272

Foucault, Michel. *Birth of the Clinic*. Translated by A.M. Sheridan Smith. Vintage Press, 1963.

———. “A Preface to Transgression.” in *Essential Work of Foucault, Volume 3,* edited by James D. Faubion. Translated by Donald F. Bouchard and Sherry Simon. New Press, 1963/1998.
———. *The Archaeology of Knowledge.* Translated by A.M Sheridan Smith. Vintage Books, 1969.
———. “Nietzsche, Genealogy, History.” In *Essential Works of Foucault Vol. 2*, edited by James D. Faubion. Translated by Donald F. Brouchard and Sherry Simon. The New Press, 1970/1998.
———. *Discipline and Punish*. Translated by Alan Sheridan. Vintage Books, 1975.
———. “The Confession of the Flesh.” In *Power/Knowledge: Selected Interviews and Other Writings, 1972-1997,* edited by Colin Gordon. Vintage Books, 1977/1980.
———. “The History of Sexuality.” In *Power/Knowledge: Selected Interviews & Other Writings 1972-1977*, edited by Colin Gordon. Translated by Alessandro Fontana and Pasquale Pasquino. Vintage Press. 1977/1980.
———. “Truth and Power.” In *Power/Knowledge: Selected Interviews & Other Writings 1972-1977*, edited by Colin Gordon. Translated by Alessandro Fontana and Pasquale Pasquino. Vintage Press: 1977/1980.
———. “What Is Enlightenment?” In *The Foucault Reader*, edited by Paul Rabinow. Translated by Catherine Porter. Pantheon Books, 1984.
———. *The Birth of Biopolitics: Lectures at the Collège de France, 1978-79.* Translated by Graham Burchell. Palgrave Macmillan, 2008.
Fraser, Nancy. “Foucault on Modern Power: Empirical Insights and Normative Confusions.” *PRAXIS International* 37, no. 1 (1981): 272-87.
Frazer, Elizabeth. “Including the Uncountable: The Realism of Feminist Politics.” *Juncture* 21, no.4 (2015): 315-318.
Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford University Press, 2007.
———. “Forgiveness—An Ordered Pluralism.” *Australasian Philosophical Review* 3, no. 3 (2019): 241-260.
———. “Scepticism and The Genealogy of Knowledge: Situating Epistemology in Time.” *Philosophical Papers* 37, no. 1 (2008): 27-50.
———. “What Is the Point of Blame?: A Paradigm Based Explanation.” *Nous* 50, no. 1 (2016): 165-183.
———. “Williams’ ‘Philosophical Anthropology’: A Humean and Nietzschean Synergy.” *European Journal of Philosophy* 24, no. (2025): 229-247.
Friedman, Marilyn. “The Practice of Partiality.” *Ethics* 101, no. 4 (1991): 818-835.
Frye, Marilyn. *The Politics of Reality: Essays in Feminist Theory*. The Crossing Press, 1983.
———. “Oppression.” In *Gender Basics: Feminist Perspectives on Women and Men*, edited by Anne Minas. Wadsworth Publishing Company, 1993.

Galston, William. “Realism in Political Theory.” *European Journal of Political Theory* 9, no.4 (2010): 385-411.

Geuss, Raymond. *Philosophy and Real Politics*. Princeton University Press, 2008.

Gordon, Colin, ed., *Power/Knowledge: Selected Interviews & Other Writings 1972-1977*. Vintage Press, 1977/1980.

Gutting, Gary. “Michel Foucault: A User’s Manual.” In *Cambridge Companion to Foucault* edited by Gary Gutting. Cambridge University Press, 1994.

Gutting, Gary. *Michel Foucault’s Archaeology of Scientific Reason*. Cambridge University Press, 1989.

Habermas, Jürgen. “The Concept of Human Dignity and the Realistic Utopia of Human Rights.” *Metaphilosophy* 41, no. 4 (2010): 464-480.

Hacking, Ian. *Historical Ontology*. Harvard University Press, 2002.

———. “Critical Notice: Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy*.” *Canadian Journal of Philosophy*, 34 no. 1, (2004): 137-148.

Hall, Edward. “Contingency, Confidence, and Liberalism in the Political Thought of Bernard Williams.” *Social Theory and Practice* 40, no. 4 (2014): 545-569.

———. *Value, Conflict, and Order: Berlin, Hampshire, Williams, and the Realist Revival in Political Theory*. Chicago University Press, 2020.

Hall, Edward and Dimitrios Tsarapatsanis. “Human Rights, Legitimacy, Political Judgement.” *Res Publica* 27 (2020): 171-185.

Han [-Pile], Béatrice. *Foucault’s Critical Project: Between the Transcendental and the Historical*. Translated by Edward Pile. Stanford University Press, 2002.

Harcourt, Bernard. “On Critical Genealogy.” *Contemporary Political Theory* 24 no.2 (2025): 167-185.

Hart, H.L.A. *The Concept of Law*. Oxford University Press, 1961.

Herodotus. *The History*. Translated by David Grene. University of Chicago Press, 1987.

Hirji, Sukaina. “Oppressive Double Binds.” *Ethics* 131, no. 4 (2021): 643-669.

Hoffman, Stanley, ed. *Political Thought and Political Thinkers*. University of Chicago Press, 1998.

Homer. *Iliad*. Translated by Richard Lattimore. University of Chicago Press, 1951.

Hoy, David. “Foucault’s Slalom.” *The London Review of Books* 4, no. 20 (1982).

Jean-Luc Goddard, dir. *Breathless (À bout de souffle*). Criterion Collection, 1960.

Jenkins, Mark P. *Bernard Williams*. McGill-Queen’s University Press, 2006.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor and Jens Timmermann. Cambridge University Press, 1785/2012.

Kateb, George. *Human Dignity.* Harvard University Press, 2001.

Khader, Serene J. *Adaptive Preferences and Women’s Empowerment*. Oxford University Press, 2001.

Killmister, Suzy. *Contours of Dignity.* Oxford University Press, 2016.

Kittay, Eva Feder. "Equality, Dignity, and Disability." In *Perspectives on Equality*, edited by Mary Ann Lyons and Fionnuala Waldron. Liffey Press, 2005.

Koopman, Colin. "Foucault's Historiographical Expansion: Adding Genealogy to Archaeology." *Journal of the Philosophy of History* 2 no. 3 (2008): 338-362.

———. "Bernard Williams on Philosophy's Need for History." *The Review of Metaphysics* 64 (2010): 3-30.

———. *Genealogy as Critique: Foucault and the Problems of Modernity*. Indiana University Press, 2013.

———. *How We Became Our Data: A Genealogy of the Informational Person*. University of Chicago Press, 2019.

———. "Genealogy, Empirical History, and Functional Teleology." *Mind*. November, 2025 https://doi.org/10.1093/mind/fzaf038.

Korsgaard, Christine. *The Sources of Normativity.* Cambridge University Press, 1996.

Kymlicka, Will. "Human Rights Without Human Supremacism." *Canadian Journal of Philosophy* 48, no. 6 (2018): 763-792.

Leeke, Wiliam D. *The Emerging Rights of the Confined*. The Correctional Development Foundation, 1972.

Longino, Helen. *Science as Social Knowledge*. Princeton University Press, 1990.

———. "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy." In *Feminism, Science, and the Philosophy of Science,* edited by L.H. Nelson and J. Nelson. Kluwer Academic Publishers, 1996.

Lorenzini, Daniele. *The Force of Truth: Critique, Genealogy, and Truth-Telling in Michel Foucault*. University of Chicago Press, 2023.

Luraghi, Nino. "Meta-historie: Method and Genre in the Histories." In *The Cambridge Companion to Herodotus*, edited by Carolyn Dewald and John Marincola. Cambridge University Press, 2006.

Macklin, Ruth. "Dignity is a Useless Concept." *BMJ* 326, no. 20-27 (2003): 1418-1420.

Manne, Kate. "Internalism About Reasons: Sad but True?" *Philosophical Studies* 167, no. 1 (2014): 89-117.

McArdle, Elaine. "The End of the Death Penalty: 'Unintended consequences' and the legacy of *Furman v. Georgia.*" *Harvard Law Bulletin* (2023).

McKittrick-Sweitzer, Lavender. "Care Exploitation," *Feminist Philosophy Quarterly* 9, no. 3 (2023).

McWhorter, Ladelle. *Racism and Sexual Oppression in Anglo-America: A Genealogy*. Indiana University Press, 2009.

———. *Unbecoming Persons: The Rise and Demise of the Modern Moral Self*/ University of Chicago Press, 2025.

Mills, Charles. *The Racial Contract*. Cornell University Press, 1997.

———. "Ideal Theory as Ideology." *Hypatia: A Journal of Feminist Philosophy* 20, no. 3 (2005): 165-183.

———. "White Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana. New York: SUNY Press, 2007.

Moore, A.W., ed. *Philosophy as a Humanistic Discipline*, edited by A.W. Moore. Princeton University Press, 2006.

Moyn, Samuel. "The Secret History of Constitutional Dignity." *Yale Human Rights & Development Law Journal* 25 (2012): 1-42.

Moyn, Samuel. *Human Rights and the Uses of History*. Yale University Press, 2014.

NAACP Legal Defense and Educational Fund. "*Furman v. Georgia*," Effective January, 2026, https://www.naacpldf.org/case-issue/furman-v-georgia/#/.

Narayan, Uma. *Upheavals of Thought: Intelligence of Emotions.* Cambridge University Press, 2001.

National Sherrif's Association. *Jail Officer's Training Manual.* National Institute of Corrections, U.S. Department of Justice, 1986.

National Sherrif's Association. *Jail Officer's Training Manual.* National Institute of Corrections, U.S. Department of Justice, 1986.

Nussbaum, Martha. "The Worth of Human Dignity: Two Tensions in Stoic Cosmopolitanism." In *Philosophy and Power in the Graeco-Roman World: Essays in Honor of Miriam Griffin*, edited by Gilian Clark and Tessa Rajak. Oxford University Press, 2002.

Nussbaum, Martha. *Frontiers of Justice: Disability, Nationality, Species Membership.* Belknap Press, 2007.

———. *Creating Capabilities*. Harvard University Press, 2011.

———. *Anger and Forgiveness: Resentment, Generosity, Justice*. Oxford University Press, 2016.

O'Neill, Onora. "Abstraction, Idealization, and Ideology in Ethics." In *Moral Philosophy and Contemporary Problems*, edited by J.D.G. Evans. Cambridge University Press, 1987.

Okin, Susan Moller. *Justice, Gender, and the Family*. Oxford University Press, 1989.

Oksala, Johanna. *Foucault on Freedom*. Cambridge University Press, 2009.

Pappas, Nick, ed., *The Jail: Its Operation and Management*. United States Bureau of Prisons, 1973.

Powers, Gershom. *A Brief Account of the Construction, Management, and Discipline of the New York State Prison at Auburn*. U.F. Doubleday, 1826.

Queloz, Matthieu. “How Genealogies Can Affect the Space of Reasons.” *Synthese* 192 (2020): 2005-2027.

———. “Revealing Social Functions Through Pragmatic Genealogies.” In *Social Functions in Philosophy: Metaphysical, Normative, and Methodological Perspectives*, edited by Rebekka Hufendiek, Daniel James, and Raphael van Riel. Routledge Press, 2020.

———. *The Practical Origins of Ideas.* Oxford University Press, 2021.

Rao, Neomi. “Three Concepts of Dignity in Constitutional Law.” *Notre Dame Law Review* 86, no. 1 (2013): 183-272.

Rawls, John. *A Theory of Justice.* Harvard University Press, 1971.

———. “Justice as Fairness: Political Not Metaphysical.” *Philosophy and Public Affairs* 14, no. 3 (1985): 223-251.

———. *Political Liberalism*. Columbia University Press, 1995.

———. *Justice as Fairness: A Restatement*. Harvard University Press, 2001.

Rossi, Enzo and Matt Sleat. “Realism in Normative Political Theory.” *Philosophy Compass* 9, no. 10 (2014): 689-701

Rouse, Joseph. “Power/Knowledge.” In *The Cambridge Companion to Foucault*, edited by Gary Gutting. Cambridge University Press, 2006.

Sangiovanni, Andrea. “Normative Political Theory: A Flight from Reality?” In *Political Thought and International Relations*, edited by Duncan Bell. Oxford University Press, 2009.

———. *Humanity Without Dignity*. Harvard University Press, 2017.

Sen, Amartya. “Gender and Co-operative Conflicts.” In *Persistent Inequalities, Women and World Development*, edited by Irene Tinker. Oxford University Press, 1990.

Shklar, Judith. *After Utopia: The Decline of Political Faith.* Princeton University Press, 1957.

———. *Legalism: Laws, Morals, and Political Trials*. Harvard University Press, 1964.

———. *Ordinary Vices.* Harvard University Press, 1984.

———. “The Liberalism of Fear.” In *Liberalism and the Moral Life*, edited by Nancy L. Rosenblum. Harvard University Press, 1989.

———. *The Faces of Injustice*. Yale University Press, 1990.

———. “The Political Theory of Utopia.” In *Political Thought and Political Thinkers*, edited by Stanley Hoffman. University of Chicago Press, 1998.

———. “What is the Use of Utopia?” In *Political Thought and Political Thinkers*, edited by Stanley Hoffman. University of Chicago Press, 1998.

Sleat, Matt. “Bernard Williams and the Possibility of a Realist Political Theory.” *European Journal of Political Theory* 9, no. 4 (2010): 485-503.

———. “Legitimacy in Realist Thought.” *Political Theory* 42, no. 3 (2014): 314-337.

Smyth, Nicholas. “Purity and Practical Reason: On Pragmatic Genealogy.” *Ergo: An Open Access Journal of Philosophy* 10, no. 37 (2023): 1057-1081.

Sophocles, *Antigone*. Translated by Robert Fagles. Penguin, 1984.
Srinivasan, Amia. "Feminism and Metaethics." In *The Routledge Handbook of Metaethics*, edited by Tristram McPherson and David Plunkett. Routledge, 2017.
———. "The Aptness of Anger." *The Journal of Political Philosophy* 26, no. 2 (2018): 123-144.
———. "Genealogy, Epistemology, and Worldmaking." *Proceedings of the Aristotelian Society* 119, no. 3 (2019): 127-156.
Steiker Carole and Jordan Steiker. *Courting Death: The Supreme Court and Capital Punishment*. Harvard University Press, 2016.
Stemplowska, Zofia. "What's Ideal About Ideal Theory." *Social Theory and Practice* 34, no. 3 (2008): 319-340.
Stemplowska, Zofia and Adam Swift. "Ideal and Nonideal Theory." In *The Oxford Handbook of Political Philosophy,* edited by David Estlund. Oxford University Press, 2013.
Strawson, P.F. "Freedom and Resentment." *Proceedings of the British Academy* 49 (1963): 187-211.
Swain, Simon. "Man and Medicine in Thucydides." *Arethusa* 27, no. 3 (1994): 303-327.
Thompson, Heather Ann. *Blood in the Water*. Pantheon Books, 2016.
Tickner, J. Ann. A *Feminist Voyage Through International Relations*. Oxford University Press, 2014.
Tiisala, Tuomo. *Power and Freedom in the Space of Reasons: Elaborating Foucault's Pragmatism*. Routledge Press, 2024.
United General Assembly. Resolution 70/175 *United Nations Standard Minimum Rules for the Treatment of Prisoners (the Nelson Mandela Rules*). 2015.
Urban Walker, Margaret. *Moral Repair: Reconstructing Moral Relations after Wrongdoing.* Cambridge University Press, 2006.
Valentini, Laura. "On the Apparent Paradox of Ideal Theory." *The Journal of Political Philosophy* 17, no. 3 (2008): 332-355.
———. "Ideal vs. Non-Ideal Theory: A Conceptual Map." *Philosophy Compass* 7, no.9 (2012): 654-664.
Veyne, Paul. *Writing History: Essay on Epistemology*. Wesleyan University Press, 1984.
Waldron, Jeremey. *Dignity, Rank, and Rights*. Oxford University Press, 2012.
Williams, Bernard. *Morality: An Introduction.* Cambridge University Press, 1972.
———. *Moral Luck*. Cambridge University Press, 1982.
———. *Ethics and the Limits of Philosophy.* Harvard University Pres, 1985.
———. "Resisting the Avalanche." *London Review of Books* 7, no. 10, 1985.
———. *Shame and Necessity*. University of California Press, 1993.
———. *Making Sense of Humanity.* Harvard University Press, 1995.
———. "On Hating and Despising Philosophy," *London Review of Books* 18, no. 8 (1996).

———. *Truth and Truthfulness: An Essay in Genealogy.* Princeton University Press, 2002.
———. *In the Beginning was the Deed.* Princeton University Press, 2005.
———. "Philosophy as a Humanistic Discipline." In *Philosophy as a Humanistic Discipline*, edited by A.W. Moore. Princeton University Press, 2006.
———. "What Might Philosophy Become." In *Philosophy as a Humanistic Discipline*, edited by A.W. Moore. Princeton University Press, 2006.
Williams, Bernard and J.J.C. Smart. *Utilitarianism: For and Against.* Cambridge University Press, 1973.
Wolf, Susan. *Meaning in Life and Why It Matters.* Princeton University Press, 2012.
Young, Iris Marion. *Justice and the Politics of Difference*. Princeton University Press, 1990
———. "Punishment, Treatment, Empowerment: Three Approaches to Policy for Pregnant Addicts." *Feminist Studies* 20, no. 1 (1994): 32-57.
———. *Inclusion and Democracy*. Oxford University Press, 2000.