

MERCY AND CRIMINAL JUSTICE

Kristen Bell

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

Chapel Hill
2010

Approved by:

Gerald Postema

Susan Wolf

Thomas Hill

Ryan Preston-Roedder

Michael Corrado

ABSTRACT

KRISTEN BELL: Mercy and Criminal Justice
(Under the direction of Gerald J. Postema)

This dissertation analyzes criminal justice from the perspective of non-ideal theory. In the first half of the dissertation, I propose a new understanding of *mercy* as a moral response to injustice within existing criminal justice systems. In the second half, I argue that certain expressions of *blame* are an injustice plaguing most criminal justice systems. In short, I am highlighting a new type of injustice and suggesting a new mode of response to injustice.

In my analysis of mercy in Part I, I distinguish between two concepts of mercy in Western political thought: negative mercy and positive mercy. To grant negative mercy is to compassionately spare someone from harsh treatment that she deserves. To grant positive mercy is to respond to someone justly when unjust social rules call for a harsher response. Following Seneca and departing from most contemporary philosophical literature, I focus on the concept of positive mercy. I argue that officials within criminal justice systems have moral reason to exercise positive mercy and that most political communities have moral reason to incorporate a general practice of positive mercy into their criminal justice systems. I argue that judges who exercise positive mercy are not impermissibly derogating from rules in service of personal feelings, but are rather serving the rule of law and fulfilling their obligation to support just institutions.

In my analysis of blame in Part II, I identify a species of blame that I call abrasive blame: the expression of attitudes meant to hurt a person because she did something wrong. The political community expresses abrasive blame toward criminal offenders through the organ of the criminal justice system. Although I argue that this abrasive blame is permissible under certain conditions, the justification is fragile at best. I argue that it is unfair for the political community to abrasively blame battered offenders and fragile offenders. I raise a red flag about abrasive blame toward these offenders; I do not argue that it is necessarily wrong to punish them. I suggest that in some cases, the exercise of positive mercy might be the political community's best response to these offenders.

ACKNOWLEDGMENTS

The committee that has supervised my dissertation has been nothing short of outstanding. Special thanks to Gerald Postema, my primary advisor, for his commitment to guide me through this intellectual journey. He has been extremely giving of his time, patience, and philosophical enthusiasm. I am especially grateful for his seminar in legal philosophy in my first year and the fact that he has challenged me from day one until day done. Thanks to Susan Wolf for her thoughtful, challenging, and constructive comments. I am grateful for the four seminars I had with her, especially the seminar on Freedom and Responsibility. Discussions and work from that seminar inform Chapter 3 of this dissertation. Thanks to Tom Hill for teaching me that the best objection is a positive proposal and to keep seeking a deeper understanding. I am grateful for conversations with Ryan Preston-Roedder about all the ideas in this dissertation. His suggestions and challenges were especially helpful in Chapters 4 and 5. Thanks also to Michael Corrado for help in the early and late stages of the process, his comments helped me develop the ideas in Chapters 1 and 2.

I have been honored to be a member of the UNC Philosophy Department, a true philosophical community. Thanks to Matt Kotzen, Bob Adams, and Jim Lesher for serving on my proposal committee. I am grateful for Jim Lesher and his wife Eleanor for their help with Seneca and Latin. Thanks also to Marilyn Adams for very insightful discussion of both positive and negative mercy in Anselm.

Thanks to Clair Morrissey for conversations, constructive suggestions, and advice at every stage in this process. Special thanks to Emily Kelahan and Elanor Taylor for an “accountability scheme” that kept my project on track. Thanks to Emily for her keen interest in the project and her useful contributions to the argument (without her friendly urge to do more history, my mention of Seneca would be limited to a footnote). Thanks to Elanor for travelling the road of graduate school alongside me. I am grateful for all the times I have learned from her and leaned on her. Thanks also to Adam Cureton, Katie Elliot, Emily Given, Benjamin Bagley, Seth Bordner, Eric Mandelbaum, and Cathay Liu for helpful comments.

I am indebted to the philosophy department and Ethics in Society program at Stanford University for paving the road that led to this dissertation. I am grateful to Tamar Schapiro for advising and to Michael Bratman for infecting me with questions about moral responsibility that I have not stopped thinking about.

The process of writing this dissertation was made much easier by the daily support, advice, and laughter of my roommate Sharif Razzaque. The fact that he woke up at 6am to work in moral support – while on vacation – deserves acknowledgement in and of itself. Special thanks to Christian Loew for believing in me and helping me “use the time.” Thanks also to him for valuable conversations and suggestions.

Thanks to Café Driade for fueling this dissertation with espresso.

I would never have embarked upon or carried out this intellectual journey without the support of my family. Thanks to my aunt for proofreading and to my grandmother for her continual interest, support, and final editing assistance. My sister Andrea has served

as the best “team dissertation” captain I could ask for. Without her genuine interest in talking to me about philosophy, I would not have found my own voice.

My deepest debt is to my parents. I could not have completed this dissertation without their love and support day in and day out. Thanks to Mum for phone calls, cards, editing assistance, and for her courageous attitude to feel fear and do it anyway. Thanks to Dad for good advice (although I’d never admit it at the time), for help in times of crisis, understanding, and for hugs at just the right time. Thanks to them both for teaching me the value of education, the drive to work hard, and the importance to think for oneself. I remember talking to my father about the first long philosophy paper I wrote. “Do you actually believe that, Kristen?” “Well, I mean, kind of.” “Sounds like you need to do some more thinking.” I ended up thinking for six more years. Thanks Dad for the drive to start and thanks Mum for the patience to see it through.

TABLE OF CONTENTS

INTRODUCTION: “WHAT IS TO BE DONE?”	1
PART I	9
CHAPTER 1: TWO CONCEPTS OF MERCY	9
SECTION I: NEGATIVE MERCY	12
SECTION II: SENECA’S MERCY	18
SECTION III: POSITIVE MERCY	31
SECTION IV: STUMBLING BLOCKS FOR NEGATIVE MERCY	40
CHAPTER 2: DEFENSE OF POSITIVE MERCY	45
SECTION I: OUTLINING A PRACTICE OF POSITIVE MERCY	47
SECTION II: LAW-ABIDING POSITIVE MERCY	51
SECTION III: OUTLAW POSITIVE MERCY	55
Rule of Law and the Case Against Positive Mercy	58
The Strategy of My Response	66
Positive Mercy in Service of Law’s Rule	70
SECTION IV: AN OBJECTION	87
CONCLUSION OF PART I:	92
PART II	95
CHAPTER 3: THE STING OF BLAME	95

SECTION I: ABRASIVE BLAME	98
SECTION II: IS ABRASIVE BLAME PERMISSIBLE?.....	105
SECTION III: REASONABLE DISAGREEMENT ON ABRASIVE BLAME.....	109
The Qualified Retributivist	110
The Hard Incompatibilist	112
The Qualified “Pacifist”	113
The Tough-Love “Pacifist”	116
An Objection.....	117
SECTION IV: ABRASIVE BLAME IN PERSONAL RELATIONSHIPS	125
SECTION V: CONTRAST TO SCANLON.....	129
CHAPTER 4: THE POLITICAL STING OF BLAME	133
SECTION I: ABRASIVE BLAME IN CRIMINAL JUSTICE SYSTEMS	135
SECTION II: THE PERMISSIBILITY OF THE POLITICAL STING	145
The Rules	147
Why Follow the Rules?.....	150
CHAPTER 5: SUBSTANTIAL WOUNDS AND THE BANDAGE OF MERCY. 158	
SECTION I: BATTERED OFFENDERS	159
SECTION II: FRAGILE OFFENDERS	174
SECTION III: WOUNDS, NOT PERMISSIBLE STINGS	182
CONCLUSION: MERCY AS A BANDAGE?	188
WHAT IS TO BE DONE FOR BATTERED AND FRAGILE OFFENDERS?	188
WHAT ABOUT MERCY?	190
REFERENCES.....	197

INTRODUCTION: “WHAT IS TO BE DONE?”

“What you can’t put right you must try to make as little wrong as possible. For things will never be perfect, until human beings are perfect – which I don’t expect them to be for quite a number of years!”

~Sir Thomas More¹

A popular approach in the philosophy of criminal punishment is to reflect on what a just system of criminal punishment would look like if it were to exist in a nearly ideal political community. This approach offers valuable insight into the conditions under which punishment can be just. The practical application of the approach, however, is limited. It calls for political communities to satisfy the conditions under which punishment can be just, but it is relatively silent on what these communities should do if these conditions remain unsatisfied. Unfortunately, these conditions do remain unsatisfied in many political communities around the world.² Most existing systems of punishment are not fully just and requisite reform is not expected to be complete in the near future.

¹ More, *Utopia*, 42.

² For example, Herbert Morris maintains that the justice in punishment is conditional on there being a legal system that imposes an equal amount of benefits and burdens upon all. Kant argues that justice in punishment is conditional on punishment being consistent with respect for offenders (and on law being in place to ensure equal liberty among people). A Rawlsian view (and other mixed theories) requires that punishment offer more security than other reasonably available alternatives if it is to be just. I doubt that these conditions are met in political communities around the world.

In the present, Lenin’s question is pressing: What is to be done? For better or worse, philosophical analysis of punishment generally leaves us without much (if any) guidance on this question. In a world where 9.8 million people are kept behind bars, many of them unjustly, this question is too pressing to ignore. The question is especially pressing in the United States. This country has the largest incarceration rate in the world, keeping over 2.3 million men and women in prisons and jails.³ Another 5.1 people are on parole or probation.⁴ In total, 7.4 million people in America (1 in 31) are under correctional supervision – more than the population of Los Angeles, Chicago, and Boston combined.⁵ These figures have been steadily climbing for decades without much sign of stopping.⁶ If any American practices of criminal punishment are unjust (as many

³ The United States has the highest incarceration rate in the world: 756 per 100,000 people in America are incarcerated. Russia and Rwanda are respectively second and third with rates of 629 per 100,000 and 604 per 100,000 people. Almost three fifths of countries (59%) have rates below 150 per 100,000. The world prison population rate is estimated at a rate of 145 per 100,000. The US has about 25% of the world’s prison population, although it has only 4% of the world’s total population. The US may also have the largest total prison population (with 2.3 million behind bars), but this statistic is disputed. Although the Chinese government reports that it has 1.5 million sentenced prisoners, non-governmental sources estimate that China imprisons at least 2.5 million people. Either way, the United States is still the world leader in per capita incarceration. Walmsley, “World Prison Population List.”

⁴ U.S. Bureau of Justice Statistics, “Correctional Populations in the United States, 2009.”

⁵ U.S. Census Bureau, “State and County QuickFacts.”

⁶ For analysis of why figures are rising, see, for example, Garland, *The Culture of Control*. For analysis of why American figures are particularly rising, see Whitman, *Harsh Justice*, and Stuntz, “The Pathological Politics of Criminal Law.”

reasonably suspect they are⁷), then there is no time like the present to ask Lenin's question: What is to be done?

In hope of shedding some light on this question, I recommend taking a slightly different approach to philosophical analysis of criminal punishment. In addition to thinking about what a system of punishment would be like if it were to exist in a nearly ideal polity, I suggest investigating existing criminal justice systems. By "criminal justice system," I mean the institutions of a political community that are charged with the task of detecting and punishing violations of criminal law—from arrest to indictment to sentencing and parole. Investigation of these systems should seek to uncover problems of justice that may be lurking within them. Although most philosophers⁸ are not particularly well positioned to critically review all the ins and outs of criminal justice systems, they are well positioned to raise red flags about injustice. They are equipped to reveal prima facie reasons why general practices and rules of criminal justice systems might, pending fuller analysis, be unjust.⁹ But philosophical analysis of criminal punishment should not stop at raising red flags within existing systems. Philosophical analysis can also shed light on what people ought to do in the short-term and long-term to address the injustice

⁷ In the words of Supreme Court Justice Anthony Kennedy: "It requires one with more expertise in the area than I possess to offer a complete analysis, but it does seem justified to say this: Our resources are misspent, our punishments too severe, our sentences too long...In too many cases mandatory minimum sentences are unwise and unjust," (Address at the American Bar Association Annual Meeting).

⁸ Jeremy Bentham and John Stuart Mill are among the chief exceptions here.

⁹ I take Joel Feinberg (in "The Expressive Function of Punishment") and Herbert Morris (in "Persons and Punishment") as paradigm examples of philosophers who investigated existing systems of criminal punishment and raised red flags.

they find in criminal justice systems. Although most philosophers¹⁰ are not well positioned to draft legislative reform bills or detailed plans for interim action, they are equipped to highlight prima facie moral reasons for and against general modes of reform and interim action.

Given this general approach to philosophical analysis of criminal punishment, my dissertation has two core parts. In the first part, I assume that existing criminal justice systems have at least some unjust rules and I investigate what should be done in light of this fact.¹¹ Unjust rules give people moral reason to act in a number of ways, for example to work for reform, to protest, and to develop interim coping strategies. I argue that unjust rules within criminal justice systems also give officials moral reason to exercise mercy within those systems. I argue not only that officials have moral reason to exercise mercy, but also that political communities have reason to incorporate a general practice of mercy into their criminal justice systems. I am not arguing that mercy should, all-things-considered, be exercised and incorporated in criminal justice systems. I am rather claiming that there is prima facie reason to put (or perhaps more accurately resurrect) mercy on the menu of options for individual and institutional action within a less-than-fully just criminal justice system.

¹⁰ Bentham and Mill are again among the chief exceptions here.

¹¹ In the second part, I raise a red flag suggesting that existing systems may have unjust rules. But my argument does not prove that existing systems do have unjust rules. Defending this assumption is beyond my scope. If the reader disagrees with it, she is welcome to read this part of the dissertation as an investigation of how people might reform or cope with unjust rules in criminal justice systems *if such unjust rules were to exist*.

My argument in the first part of the dissertation is about how officials might use mercy to respond to injustice in systems of criminal justice. My argument in the second part raises a red flag about those systems of criminal justice. I present reason to suspect that modern criminal justice systems, especially the American system, may be harboring unjust practices and rules that have thus far escaped critical radar. These practices concern the institutional expression of what I call abrasive blame toward offenders. I identify abrasive blame as the expression of attitudes intended to sting a person for having done something wrong. I argue that although the expression of abrasive blame in the criminal justice system may be permissible toward some offenders, it is prima facie impermissible toward (at least) two types of offenders. Criminal justice systems generally have rules that call for abrasively blaming these offenders even though doing so is prima facie impermissible. I raise a red flag about these rules and I suggest that a political community has reason to reform them.

In the conclusion, I discuss how the two parts of the dissertation can work together. In the first part I argue that there is reason to exercise mercy in response to unjust rules. Then I identify a set of potentially unjust rules that could give occasion for such exercise. Put in a different order, I identify a set of potentially unjust rules and argue that insofar as they are in fact unjust, officials in the criminal justice system have reason to exercise mercy in light of them. Although the two parts of the dissertation work together, they do not strictly depend upon one another. Rejection of one part does not preclude or jeopardize acceptance of the other part. Thinking about how to address problems (Part 1) is different from thinking about which problems we have to address (Part 2). But both must be done if we are to answer the question: What is to be done?

With this general frame of the argument in place, I give a more specific outline of the chapters to follow.

Part I: A Case for Mercy

Chapter 1: Two Concepts of Mercy

In this chapter, I analyze the conceptual nature of mercy. I argue that there are two concepts of mercy: negative mercy and positive mercy. Negative mercy is compassionately sparing someone from some harsh treatment that she deserves. Positive mercy, which is not the focus of contemporary literature but which is inspired by Seneca's "On Mercy," is treating a person justly when unjust social rules call for giving her harsher treatment. I suggest that the conceptual contours of negative mercy pose significant challenges for using it in the context of most criminal justice systems. Positive mercy lacks these problematic contours and lies ready for further investigation.

Chapter 2: Defense of Positive Mercy

Having explained the concept of positive mercy in the previous chapter, I investigate its potential merit in the context of criminal justice. I distinguish between law-abiding positive mercy (which requires an official to derogate from a social rule but not a law) and outlaw positive mercy (which requires an official to derogate from a law). I argue that individual actors in criminal justice systems have reason to exercise both kinds of positive mercy. I also argue that political communities have reason to incorporate a general practice of positive mercy into their criminal justice systems. Exercises of positive mercy not only achieve substantive justice, but also help serve the rule of law.

Part II: Abrasive Blame and its Wounds

Chapter 3: The Sting of Blame

In this chapter, I identify a phenomenon that I call abrasive blame: the expression of attitudes that are meant to sting a person because she did something wrong. Abrasive blame is a common phenomenon in many people's lives and it is very common within existing criminal justice systems. It is tempting to presume that abrasive blame is a morally permissible response to any agent who freely and responsibly commits a wrong action. I challenge this presumption. Abrasive blame is only permissible when it is properly embedded in reasonable social practices that are developed within personal relationships and communities. My claim is not that abrasive blame is usually impermissible, but rather that its permissibility is conditional on the features of the relationships and communities in which it is expressed.

Chapter 4: The Political Sting

Having identified abrasive blame and the general conditions of its permissibility, I consider whether it is permissible within the context of a political community. I do so with an eye to considering whether it is permissible within the more particular context of political communities' criminal justice systems. When a political community has developed decent social rules about abrasive blame (as almost all have), there is presumptive reason for people to act in accord with such rules and engage in abrasive blame. Given this presumptive reason, many (but not all) acts of abrasive blame within political communities are *prima facie* permissible.

Chapter 5: The Sting's Wounds in Criminal Justice

Armed with the work of Chapters 3 and 4, I direct my focus to flagging injustice in existing criminal justice systems. I highlight a *general deficiency* in social rules that govern abrasive blame toward criminal offenders in several political communities. In

several of these communities (e.g. America), the rules deem it appropriate for the criminal justice system to abrasively blame what I call battered and fragile offenders. I argue that it is unfair for the political community to abrasively blame battered and fragile offenders. Criminal justice officials have strong reason not to abrasively blame these offenders and to change social rules that would license such blame. In general, blaming battered and fragile offenders usually does not sting them permissibly; it unfairly strikes at them and inflicts substantial wounds.

PART I.
CHAPTER 1: TWO CONCEPTS OF MERCY

“Everybody's cryin’ ‘Mercy’ When they don't know the meanin’ of the word.”
~ Mose Allison¹²

Modern political communities aim for justice under law in responding to crime. The political institutions that respond to crime are criminal *justice* systems. People expect to “see justice done” in courts of law. One may doubt that criminal justice systems achieve justice, but justice under law is certainly their aim. Justice, however, has not always been the aim of political communities’ responses to crime. At the turn of the first millennium, Seneca maintained that *mercy* is the proper aim of a political ruler’s response to crime.¹³ His letter on how rulers ought to respond to crime is not titled, “On Justice,” but “On Mercy.”¹⁴ He judges the quality of a ruler by tallying up his merciful actions. According to Seneca, a ruler should respond to all offenders with mercy. To a modern audience, this claim is bizarre in at least two ways. First, most members of a

¹² Quoted by Rainbolt, “Mercy: In Defense of Caprice,” 226.

¹³ Seneca wrote “On Mercy” in 55-56 AD. For historical context, see *De Clementia*, ed. Braund.

¹⁴ The Latin title of Seneca’s work is “De Clementia.” “Clementia” is translated as either “mercy” (see *Moral Essays*, trans. Basore and *Dialogues and Essays*, trans. Davie) or “clemency” (see *De Clementia*, trans. Braund and *Anger, Mercy, Revenge*, trans. Robert A. Kaster and Martha C. Nussbaum). Throughout the dissertation, I use “mercy” rather than “clemency.”

modern audience think that if a ruler were to respond to all offenders in one way, it would be with justice, not mercy. Second, Seneca suggests that a ruler should never withhold mercy from an offender. This position conflicts with the modern thought that mercy is supererogatory, certainly not morally required toward any offender, let alone every offender. Why is Seneca's discussion of mercy so jarring to the modern ear? Should we bend our ear to listen to his discussion or reject it as out of tune?

Loosely echoing the strategy of Isaiah Berlin in "Two Concepts of Liberty," I identify two concepts of mercy in Western political thought: negative mercy and positive mercy. To grant negative mercy is to compassionately *spare* someone *from* harsh treatment (generally in the form of punishment) that she deserves. To show positive mercy is to *respond to* someone justly when unjust social rules call for a harsher response (generally in the form of punishment). Contemporary people generally use the term "mercy" to mean negative mercy whereas Seneca and other pre-modern thinkers generally use it to mean something closely akin to positive mercy. The stark contrast between these two concepts explains why Seneca's discussion jars the modern ear.

The purpose of this and the following chapter, however, is not to track changes in the use of the term "mercy" over time. My aim is to compare negative mercy and positive mercy with an eye to investigating whether there is good reason to exercise either (or both) in a criminal justice system. I argue that although negative mercy faces significant challenges for use in most criminal justice systems, positive mercy does not face such challenges. In the next chapter, I explore positive mercy more fully and argue that there is reason to exercise it in most criminal justice systems. My argument is not that positive mercy is closer to some metaphysical truth about Mercy (I doubt there is

one), but that it is likely more apt for use in most criminal justice systems. We should not only bend our ear to Seneca's discussion of mercy, but we should also maintain, or perhaps resurrect, something like his concept of positive mercy in our working repertoire of political and jurisprudential concepts.

The structure of the chapter is as follows. In Section I, I present the concept of negative mercy which is at the forefront of most contemporary philosophical and jurisprudential literature on mercy. Negative mercy is compassionately sparing someone from deserved harsh treatment (generally in the form of punishment). In Section II and III, I present the concept of positive mercy. I begin by offering an interpretation of Seneca's "On Mercy" in Section II. My interpretation is intended to be charitable and true to his work, but my main objective is not to pin down historic Seneca. Instead, I am climbing on what I take to be his shoulders in order to see and spell out what I call positive mercy. In Section III, I identify positive mercy as treating someone justly when unjust social rules call for harsher treatment (generally in the form of punishment). I close this section with a comparison of positive mercy and negative mercy. They are both the prerogative of a person with power to give or withhold harsh treatment, but they share little else in common.

After having identified and contrasted negative and positive mercy, I turn to analyze their respective merit in the context of criminal justice. In Section III, I argue that negative mercy faces significant challenges for use in a criminal justice system. The features that count as strikes against negative mercy in the context of criminal justice do not apply to positive mercy. In the next chapter, I show that there is reason to exercise positive mercy in most criminal justice systems.

Section I: Negative Mercy

Most philosophers and lawyers writing about mercy over the past fifty years share a general concept of mercy that I call *negative mercy*. Negative mercy is sparing someone from deserved suffering (generally the suffering of punishment) out of compassion or some other special concern for her. Negative mercy is distinct from justice, it is supererogatory, motivated by compassion or some other form of special concern, and unrestrained by reason. In this section, I consider each of these features in turn and show that it is indeed part of the received view on mercy in contemporary political thought. Most authors endorse most of these features, but some reject one or more. I found no author, however, who rejects any features of the received view without argument.¹⁵ For the purpose of this section, I will use the term “mercy” to mean negative mercy.

¹⁵ The following is a reasonably comprehensive list of contemporary publications about mercy: Moore, *Pardons: Justice, Mercy, and the Public Interest*; Murphy and Hampton, *Forgiveness and Mercy*; Sarat and Hussain, *Forgiveness, Mercy, and Clemency*; Adler, “Murphy and Mercy;” Bennett, “The Limits of Mercy;” Brett, “Mercy and Criminal Justice;” Brien, “Mercy, Utilitarianism, and Retributivism,” “Mercy and Desert,” and “Mercy within Legal Justice;” Card, “On Mercy;” Cartwright, “Revenge, Punishment, and Mercy;” Dolinko, “Some Naïve Thoughts about Justice and Mercy;” Duff, “The Intrusion of Mercy;” Garvey, ““As the Gentle Rain from Heaven;” Harrison, “The Equality of Mercy;” Hestevold, “Justice to Mercy;” Hurd, “The Morality of Mercy;” Johnson, “Entitled to Clemency;” Kleinig, “Mercy and Justice;” Kobil, “Mercy in Clemency Decisions;” Markel, “Against Mercy;” Nussbaum, “Equity and Mercy;” Nuyen, “Straining the Quality of Mercy;” O’Driscoll, “The Quality of Mercy;” Pearn, “The Quiddity of Mercy;” Rainbolt, “Mercy: An independent, imperfect virtue” and “Mercy: In defense of caprice;” Ransome, “Above the Sceptred Sway;” Roberts, “Mercy;” Sigler, “The Story of Justice;” Smart, “Mercy;” Statman, “Doing Without Mercy;” Steiker, “Tempering or Tampering;” Sterba, “Can a person deserve mercy?;” Tasioulas, “Mercy;” Twambley, “Mercy and Forgiveness;” Walker, “The Quiddity of Mercy.”

Contemporary philosophers and lawyers tend to define mercy as giving an offender less punishment than she deserves.¹⁶ John Tasioulas defines mercy as a “form of charity towards wrongdoers that justifies punishing them less severely than they deserve according to justice;”¹⁷ Heidi Hurd as “a properly-motivated suspension of just deserts;”¹⁸ H. Scott Hestevold as “tempering deserved suffering,”¹⁹ and Daniel T. Kobil as “an act of benevolence or compassion that reduces what [punishment] is owed.”²⁰ Others, like Jeffrie Murphy and Jean Hampton, have slightly different definitions of mercy, but they agree that mercy necessarily involves giving less than the deserved amount of punishment. Peter Geach agrees that mercy must involve giving an offender less than he deserves; “relaxing the penalty of a general law in a ‘deserving case’ is not mercy at all but mere justice.”²¹

¹⁶ The most obvious exceptions are Nigel Walker and George Rainbolt. Rainbolt explicitly rejects the trend in defining mercy as sparing deserved suffering; “no analysis which defines mercy in terms of the relief of deserved suffering can account for all cases of the virtue,” (“Mercy: An Independent, Imperfect Virtue,” 228). He points out that Nazis, pirates, and black knights can show mercy. Walker argues that mercy’s “nature and function are uninterestingly obvious. It merely allows benign interference when the programming of the system seems to be having unacceptable effects in special cases,” (“Quiddity of Mercy,” 27). Walker, like Seneca but unlike the vast majority of contemporary writers about mercy, seems to endorse something quite close to the concept of positive mercy.

¹⁷ “Mercy,” 101.

¹⁸ “The Morality of Mercy,” 389.

¹⁹ “Justice to Mercy,” 281.

²⁰ “Mercy in Clemency Decisions,” 39.

²¹ *Truth and Hope*, 96.

Mercy is not a part of justice, but rather a departure from justice; Tasioulas writes that mercy and justice are “incommensurable values” and that mercy has an “irreducible distinctiveness” from justice.²² Mercy is “an autonomous virtue, not reducible to justice,” according to Murphy²³ and mercy is an “independent virtue...[that] modifies the demands of justice” according to George Rainbolt.²⁴ Ross Harrison takes it as axiomatic that mercy is not part of justice.²⁵ ²⁶

Mercy is not morally required, but rather morally optional.²⁷ Many contemporary authors take themselves to be capturing the thought expressed in Shakespeare’s line, “the quality of mercy is not strained.”²⁸ Carol Steiker describes mercy as “irreducibly supererogatory.”²⁹ Murphy claims that “mercy transcends the realm of strict moral obligation and is best viewed as a free gift; an act of grace, love, or compassion that is

²² “Mercy,” 108, 114.

²³ *Forgiveness and Mercy*, 169.

²⁴ “Mercy: An Independent, Imperfect Virtue,” 169.

²⁵ “The Equality of Mercy.”

²⁶ Exceptions to this view include: Kleinig, Johnson, and perhaps Brien.

²⁷ Exceptions to this view include: Smart, Tasioulas, Walker, and perhaps Card. Card writes, “it is difficult to recognize an action as the showing of mercy if the agent was obligated to act as he did anyway,” but she thinks mercy is morally deserved and wrongfully withheld under certain circumstances (“On Mercy,” 184). Although Tasioulas denies that mercy is supererogatory, he recognizes that he is arguing against the grain in this vein; “it is widely taken as axiomatic that granting mercy is a supererogatory act,” (“Mercy,” 125).

²⁸ *The Merchant of Venice*, Act IV, Scene 1, line 179–197.

²⁹ “Tempering or Tampering,” 26.

beyond the claims of right, duty, and obligation.”³⁰ Hampton argues that mercy is always a gift because the alternative course of action (giving retributively just deserts) is always a morally legitimate option.³¹ Hestevold writes, “Mercy involves the sparing of deserved suffering when there is no moral obligation to do so.”³² According to A.T. Nuyen, mercy is “conceptually exclusive of rights and obligations;” like grace, mercy is “free and unmerited... If a judge ought to give the offender a certain treatment or if he is obligated to give the offender a certain treatment, then that treatment cannot be merciful. By the definition of mercy, one can never be obligated to show mercy.”³³ P. Twambley concurs, maintaining that mercy is freely given, “it is not something to which a man is compelled by the claim of other obligations.”³⁴

Although acts of mercy are not morally required, it is nevertheless morally good or virtuous for a person to perform them. Mercy is taken to be an important virtue and merciful actions are generally morally good. Murphy writes that “mercy is an autonomous virtue” and “a moral virtue.”³⁵ Rainbolt maintains that as a matter of conceptual necessity, mercy is “prima facie morally good.”³⁶ Antony Duff concurs,

³⁰ *Forgiveness and Mercy*, 166.

³¹ *Ibid.*, 159-161.

³² “Justice to Mercy,” 288.

³³ “Straining the Quality of Mercy,” 66.

³⁴ “Mercy and Forgiveness.”

³⁵ *Forgiveness and Mercy*, 166.

³⁶ “Mercy: In Defense of Caprice,” 234.

holding that mercy is right and virtuous by definition; “mercy is, of course, a virtue, and to act mercifully is to act rightly.”^{37 38}

Almost all contemporary philosophers agree that mercy must be motivated by sympathy, pity, or some other emotional concern for its recipient.³⁹ As Hampton puts it, mercy must be “granted out of pity and compassion for the wrongdoer.”⁴⁰ Murphy claims that people have a strong intuition that mercy involves “a character disposed to perform merciful acts from love or compassion.”⁴¹ Claudia Card, Tasioulas, and Duff all agree that mercy must be motivated by special concern for the recipient. According to Jacob Adler, “a person possessing the virtue of mercy must be able to feel compassion, to recognize these feelings, and to act on them; and be disposed to act on them.”⁴² Nuyen extends the point a bit further and holds that the very function of “mercy is to benefit oneself by relieving a psychological discomfort.”⁴³

³⁷ “The Intrusion of Mercy,” 361.

³⁸ Among others, Steiker and Bennett disagree with this point. They each maintain that mercy can be either good or bad depending on its use.

³⁹ Rainbolt is an exception. He recognizes cases of “cold-hearted” mercy such as a Nazi showing mercy without compassion. “Mercy: An Independent, Imperfect Virtue,” 170-171.

⁴⁰ *Forgiveness and Mercy*, 158.

⁴¹ *Ibid.*, 166.

⁴² “Murphy and Mercy,” 262-263.

⁴³ “Straining the Quality of Mercy,” 72. Nuyen suggests that Seneca agrees with her idea that one shows mercy in order to relieve psychological discomfort at the sight of another’s suffering (*ibid.*, 70-71). I disagree. As I will explain in the next section, Seneca thinks one should not feel pity or any kind of psychological discomfort in the first place. He is a Stoic and rejects pity as the “defect of a small mind.” “On Mercy,” in *Dialogues and Essays*, 215.

Most also hold that mercy floats free from reason. A person may have motivational or justificatory reasons for showing mercy, but she need not.⁴⁴ Writing on the closely related concept of clemency, the U.S. Court of Appeals for the Sixth Circuit holds “the very nature of clemency is that it is grounded solely in the will of the dispenser of clemency. He need give no reasons for granting it, or for denying it.”⁴⁵ Mercy is not necessarily motivated by or justified by reasons. Some authors make an even stronger claim that mercy necessarily cannot be justified by reasons. Harrison leads the charge here. Riffing off Shakespeare, Harrison maintains that if mercy is “[con]strained by reasons” or “given for reasons,” then it is not mercy.⁴⁶ For Harrison, mercy is necessarily an arbitrary act; “reasons squeezes out mercy.”⁴⁷ H.R.T Roberts seems to agree, maintaining that mercy cannot be justified by reasons; “a genuine act of mercy is always unjustified.”⁴⁸ Most modern authors prefer the weak claim that reasons do not necessarily motivate or justify mercy. Only a handful of authors endorse the strong claim that justificatory reasons are necessarily absent in grants of mercy. Although only some think that mercy is allergic to reasons, most agree that mercy is not governed by reasons.

Current scholarship about mercy is predominantly focused upon the concept of negative mercy. As I have explained in this section, negative mercy is giving a person

⁴⁴ Tasioulas is somewhat of an exception. Mercy need not be given for reasons, but the only defensible cases of mercy are justified by reasons.

⁴⁵ *McQueen v. Paton*, 118 F.3d 460, 465 (6th Cir. 1997).

⁴⁶ “The Equality of Mercy,” 109.

⁴⁷ *Ibid.*, 110.

⁴⁸ “Mercy,” 353.

less harsh treatment, generally punishment, than she deserves. It is distinct from justice, morally supererogatory, driven by compassion or pity (or some other emotion), and not necessarily governed by reasons. As I show in Section III, the concept of negative mercy runs into trouble quickly in the context of most criminal justice systems. Before discussing this trouble, however, I want to contrast the concept of negative mercy in contemporary political thought with the concept of positive mercy that has roots in Seneca's work.

Section II: Seneca's Mercy

Contemporary philosophers writing on mercy claim that Seneca's "De Clementia" ("On Mercy") is the first "self-conscious" discussion of the idea of mercy.⁴⁹ I argue, however, that Seneca is not focused on the same concept of mercy that these commentators generally employ.⁵⁰ Seneca states that mercy is not properly understood as "the moderation that removes something from the due and merited punishment."⁵¹

⁴⁹ See Tasioulas, Nussbaum, and Steiker. Cicero's extensive discussion of and appeal for "misericordia" in his speech defending a young politician appears to have escaped the attention of these commentators. See *Pro Murena*, *Pro Marcello*, *Pro Ligario*, and *Pro Rege Deiotaro*. Ironically, Cicero appeals to the concept of negative mercy that contemporary philosophers have their eye on whereas Seneca does not. As I argue, Seneca's work focuses on the different concept of positive mercy. See Braund, *De Clementia* for a careful history of "clementia" in Roman history.

⁵⁰ For an interpretation of Seneca that echoes my own, see Barrozo, "Punishing Cruelly." Barrozo maintains that Seneca did not just advocate for leniency in specific cases, but rather pushed for overall penal reform on grounds that the Roman system was cruel. "On Mercy," according to Barrozo's interpretation is, "a systematic reflection on the repudiation of cruelty and the corresponding duty of mercy in the context of punishment," ("Punishing Cruelly," 71). My interpretation of Seneca is different from, but not strictly inconsistent with Martha Nussbaum's in "Equity and Mercy." (Personal conversation with Nussbaum, April 16, 2010).

⁵¹ "On Mercy," in *Dialogues and Essays* trans. Davie, 214.

Seneca admits that according to popular opinion, “mercy consists in stopping short of the penalty that might have been deservedly fixed.”⁵² He argues, however, that popular opinion is wrong on this point (indeed the next sentence begins, “men of limited judgment think...”⁵³).⁵⁴ He goes on to say that *pardon*, not *mercy*, is giving an offender less punishment than she deserves; “pardon is the cancellation of punishment that is due.”⁵⁵ Seneca argues that the truly wise man does not pardon. The argument is simple: “pardon is extended to a man who ought to be punished; but a wise man does nothing which he ought not to do, and leaves undone nothing which he ought to do; accordingly he does not cancel a punishment which he ought to exact.”⁵⁶ Seneca is very clear that mercy is different from pardon; granting mercy is virtuous, pardoning is vicious. As he puts it, “the superiority of mercy [over pardon] lies primarily in this, that it declares that

⁵² Ibid.

⁵³ Ibid.

⁵⁴ It was a common strategy to consider popular opinion and sometimes reject it as wrong. The strategy takes its cue from Aristotle’s method of considering popular opinions as *endoxa* (significant but rejectable data points). A close look at the original Latin text suggests a slightly different interpretation. Seneca writes, “Atqui hoc omnes intellegunt clementiam esse, quae se flectit citra id, *quod merito consisui posset*,” (emphasis mine); everyone understands the fact that mercy consists in stopping short of *what is established as merited* (emphasis mine). Is Seneca reporting the popular opinion that mercy is giving less than what society *deems as* deserved (which is not necessarily less than what is *actually* deserved)? If so, Seneca does not think popular opinion is wrong, but simply in need of careful interpretation. I am indebted to Jim and Eleanor Leshner for help with the Latin translation.

⁵⁵ “On Mercy,” in *Dialogues and Essays* trans. Davie, 217.

⁵⁶ Ibid.

those who escape punishment should not have been treated in any way differently; it is more rounded than pardon, and more honorable” (emphasis mine).⁵⁷

To someone focused on the concept of negative mercy, Seneca’s praise for mercy appears blatantly inconsistent with his opposition to giving offenders less than they deserve. Equally odd is Seneca’s harsh criticism of pity that moves people to give offenders less than they deserve; such pathos is the “defect of a small mind”⁵⁸ according to Seneca the Stoic. Surely Seneca is not using the concept of negative mercy that is popular today. In what follows, I argue that Seneca is discussing and defending a different sort of mercy that is closely akin to what I call *positive mercy*. My project in this section is to give an interpretation of Seneca’s work on mercy. In the next section, I draw heavily on his work to spell out the concept that I call positive mercy.

My interpretation begins with Seneca’s own definition of mercy. For Seneca, mercy is a virtue⁵⁹ and it is “the inclination of the mind towards mildness in exacting punishment.”⁶⁰ This remark is difficult to interpret because Seneca does not indicate the standard according to which mildness counts as mildness. Mildness is a relative term; what was mild in punishment under the reign of Draco (cutting off one finger instead of the whole hand) is certainly not mild today. Mildness in punishment is giving less harsh punishment, but less harsh according to what standard? For Seneca, mildness cannot

⁵⁷ Ibid., 218.

⁵⁸ Ibid., 215

⁵⁹ He often uses the term, as I generally do, to refer to a type of action (namely the type of action that someone with the virtue of mercy is inclined to engage in).

⁶⁰ “On Mercy,” in *Dialogues and Essays* trans. Davie, 214.

mean giving less than the deserved punishment, for he opposes giving less than deserved punishment. I interpret him to mean that mildness in punishment is *giving less punishment than the amount prescribed by operative social rules*.⁶¹

Seneca does not explicitly mention any social rules, let alone social rules that prescribe punishment. By “social rules,” I mean the existing laws as well as the predominant cultural norms, customs, and social mores.⁶² The punishment prescribed by the operative social rules is the expected or standard punishment in a given culture or legal system. At Seneca’s time (1BC - 65 AD), crucifixion⁶³ and being whipped to death were standard punishments for treason, enslavement was a standard punishment for theft, and banishment was a standard punishment for forgery. Criminals were often condemned to serve as gladiators (*venatores*), to be burned alive (*crematio*), or to be fed to bears, lions, or panthers (*damnatio ad bestias*) for public entertainment. Many criminals were sentenced to play the roles of characters in gruesome Greek and Roman myths, which involved being raped, castrated, and/or being eaten by beasts in the

⁶¹ Seneca also writes that mercy is the “forbearance of a superior towards an inferior in determining punishment,” (ibid.). Again, interpretation is difficult here because forbearance is a relative term. “Forbearance” does not tell us much unless we know what the merciful superior forbears from. The truly merciful superior certainly does not forbear from giving deserved punishment. Instead, I suggest that the merciful superior forbears from giving an inferior the full amount of punishment suggested by the social rules.

⁶² I fill out a more detailed account of social rules in “The Political Sting,” but the brief description is all that is needed here.

⁶³ In his *Dialogues*, Seneca describes a particularly gruesome mode of crucifixion “whereby the victim was impaled through the genitals,” (Coleman, “Fatal Charades,” 61).

amphitheater.⁶⁴ Seneca reluctantly witnessed such punitive entertainment and writes about it with reproach in his Epistles.⁶⁵ Roman masters were free to punish slaves in any way they pleased; whipping, scourging, breaking bones, branding, and crucifixion were common.⁶⁶ Seneca criticizes the cruelty of a master who planned to throw his slave to the giant lampreys in his fishpond for having broken a dinner glass.⁶⁷ The standard punishment for parricide was to be tied in a sack with a serpent and a monkey and thrown into the sea or a river (*poena cuelli*).⁶⁸ Seneca disparages this punishment in “On Mercy,” arguing that its inclusion in the law did not reduce instances of parricide, but spurred them on; “children were shown the way to the deed by its punishment.”⁶⁹ Many of these punishments, including the *poena cuelli*, were prescribed by Roman law. Others

⁶⁴ See Coleman, “Fatal Charades.” As part of some of the most gruesome punishments, female convicts were forced to play roles that involved being raped by a bull in the amphitheater.

⁶⁵ “I happened to go to one of the lunchtime interludes, expecting there to be some light and witty entertainment, some respite for the purpose of relieving people’s eyes of the sight of human blood: far from it... In the morning men are thrown to the lions and the bears: but it is to the spectators that they are thrown in the lunch hour,” (Seneca, Epistles 7, 3-4, as quoted in Coleman, “Fatal Charades”).

⁶⁶ Coleman, “Fatal Charades.”

⁶⁷ “On Anger,” in *Dialogues and Essays* trans. Davie. According to Braund, “On Anger” and “On Mercy” are complementary texts (for her discussion of the relationship between the two texts see *De Clementia*, 70-73).

⁶⁸ By the time of Constantine, a cock and dog were also put in the sack. See Egmond, “The Cock, the Dog, the Serpent, and the Monkey” and Radin, “The Lex Pompeia and the Poena Cullei.”

⁶⁹ “On Mercy,” in *Dialogues and Essays* trans. Davie, 209.

were prescribed by less formal social norms; for example, whipping was generally prescribed as the informal punishment for a personal insult.⁷⁰

Seneca criticized the popular norms about punishment on the grounds that they were cruel and tainted by the vice of anger and revenge.⁷¹ As a Stoic, Seneca deemed anger an unconditional evil, an “unbridled and deranged madness”⁷² to which people are prone unless they are strictly tutored to control themselves. Shaped by people’s anger and their angry deeds, the social rules of his time often (if not always) prescribed cruel revenge as punishment. Seneca describes typical Roman punishments as anger’s “equipment.” In “On Anger,” he harshly criticizes Persian and Macedonian punishment practices as nothing but “savagery shown in anger”⁷³ and then bemoans the fact that customary Roman punishments are no different; “How I wish that examples of such savagery had been confined to foreign races, and that the barbarism of torture and anger had not been imported into our Roman customs, together with other vicious practices

⁷⁰ I doubt that these Roman social rules prescribed punishments that were “proportionate” to crimes. But if they were, then Nussbaum is correct in her interpretation of Seneca (see “Equity and Mercy”). According to Nussbaum, Seneca sees mercy as giving less punishment than proportionate to the crime. According to my interpretation, Seneca sees mercy as giving less punishment than prescribed by social rules. Our interpretations are different, but not strictly incompatible.

⁷¹ Seneca’s criticism of Roman punishment was not a popular opinion among those with the power to punish. The social norm at the time was to embrace and even increase cruelty in punishment. Nero did not take Seneca’s advice to show mercy and neither did Galba, the emperor for a short period after Nero. When offenders sentenced to crucifixion appealed to Galba for mercy, he responded by forcing them to carry a much larger cross that was painted white. See Braund, *De Clementia*.

⁷² “On Anger,” in *Dialogues and Essays* trans. Davie, 20.

⁷³ *Ibid.*, 33.

from abroad!”⁷⁴ He lists customary Roman punishments as the “equipment” of anger, not of justice.

[T]he equipment that is [anger’s] very own – the horse of torture, the cord, the gaol, the cross, the fires that encircle live bodies buried in the ground, the hook that drags along corpses as well, the different kinds of chains and of punishments, the tearing of limbs, branding of the forehead, the pits where monstrous beasts prowl: let anger be set in the midst of these implements, uttering a terrible and horrible shriek, more loathsome than all these instruments that let it vent its fury.⁷⁵

Seneca emphasizes a gap between the amount of punishment that justice called for and the amount of punishment that the corrupt social rules of his time called for. Seneca praises a merciful person as someone with the presence of mind (and the critical eye) to recognize this gap and to punish justly rather than cruelly. In being merciful, a person gives an offender less than the socially prescribed amount of punishment, but no more or less than the offender ought to be given. A ruler has the power and social license to inflict the punishment that social rules prescribe – cruel revenge in Seneca’s time – but the sage ruler exercises control and judges carefully to punish an offender justly. As Seneca puts it, “mercy consists in controlling the mind when one has the power to take revenge.”⁷⁶ The opposite⁷⁷ of mercy is not strictness (indeed “strictness is her natural counterpart”⁷⁸) but cruelty.⁷⁹

⁷⁴ Ibid.

⁷⁵ Ibid., 20.

⁷⁶ “On Mercy,” in *Dialogues and Essays* trans. Davie, 214.

⁷⁷ I read Seneca as understanding mercy and strictness (both virtues) in the following relationship with cruelty and pity (both vices). There are two ways for social rules not to track just punishment. The rules can demand too much punishment; they can be too harsh (as they were in Seneca’s Rome). Or the rules can demand too little punishment; they can be too lax. Cruelty is the vice that drives us to give too much punishment and

Mercy is the virtue of exercising control and treating offenders justly rather than recklessly following social rules that prescribe overly harsh, cruel punishment. In order to be merciful, one must know how to treat an offender justly. According to Seneca, an offender does not necessarily deserve to be treated in a way that is proportionate to her crime, or to be punished in a way that somehow corresponds to the gravity of her crime. Rather, three considerations guide the way a ruler ought to treat to an offender: the treatment should morally reform the offender, morally educate the general population, and decrease crime in the polity.⁸⁰ Seneca argues that the most effective way to achieve these three aims is to inflict as little harsh treatment as possible on offenders (often none at all) without blurring the public distinction between the virtuous and the vicious. On pain of slipping into viciousness and instability, society needs to hear the message that vicious people are different and are treated differently from good or reformable men. The bad men are marked with punishment, the good are let go, and the reformable are

pity is the vice that drives us to give too little punishment. Strictness is the corrective for pity that brings punishment up to the just level when social rules are too lax. Mercy is the corrective for cruelty that brings harsh punishment down to the just level when social rules are too harsh. Mercy and strictness are virtues that are “in natural harmony” because they aim at the same mean—just punishment. Full defense of this interpretation is beyond my scope because it would involve more thorough analysis of stoic theory of virtue. I am not, at the end of the day, interested in mercy as a virtue, but rather as a type of action.

⁷⁸ “On Mercy,” in *Dialogues and Essays* trans. Davie, 215.

⁷⁹ “Men of limited judgment regard strictness as the opposite of mercy; but no virtue has a virtue as its opposite. What, then, is set in opposition to mercy? It is cruelty, which is nothing other than harshness of the mind in claiming punishment,” (ibid.).

⁸⁰ Note the distinct lack of retributive aim in punishment; contemporary philosophers who think mercy only makes sense in a retributive framework clearly do not have Seneca’s idea of mercy in mind.

reformed. “Once the distinction between bad men and good is removed, what follows is confusion and the outbreak of vice; accordingly, a wise restraint should be shown, such as is capable of distinguishing between curable characters and ones past hope.”⁸¹ The ruler ought to punish offenders enough to maintain a public distinction between good and bad men, but not any more than this. Punishing offenders more than this (as Seneca’s contemporaries often did) does not improve offenders, morally educate the population, and/or reduce crime.⁸² It is not just; it is simply a pointless infliction of harsh treatment; it is cruel. According to Seneca, a person ought to be punished in a way that inflicts as little harm as possible while maintaining the public distinction between virtuous and vicious people.

Seneca claims that mercy is not indiscriminate (“the mercy we exercise ought not to be indiscriminate”⁸³), but he also holds that a ruler should show mercy to all men, good and bad; “no one is deprived of favor [mercy] from me, *though he lacks everything but the name of man*” (emphasis mine).⁸⁴ When social rules prescribe punishment above and

⁸¹ “On Mercy,” in *Dialogues and Essays* trans. Davie, 190.

⁸² Seneca gives examples of how ineffective contemporary punishments were at achieving the three aims of punishment. “A sparing hand in applying punishments is more successful in improving morals”—of both the offender and the population at large (ibid., 209). Infamy (the punishment that denies a citizen the right to speak in public forums), enslavement, and other punishments that destroy a person’s reputation are ineffective at reform because “no one shows regard for a reputation he has lost; when he no longer has any room for punishment, a man enjoys a kind of immunity from it,” (ibid.). Poena cuelli (the punishment for parricide that involves tying the offender in a sack and throwing him into the seas) is ineffective at reducing crime; its use increased rather than decreased instances of parricide in Rome. Seneca exaggerates the point, writing that “the act of killing a parent began with the law condemning it,” (ibid.).

⁸³ Ibid., 190.

⁸⁴ Ibid., 188.

beyond the amount needed to maintain the distinction between good men from bad men, a ruler should not follow these social rules. Instead, a virtuous ruler gives just as much punishment that is needed to mark the distinction between the men. He does not treat offenders indiscriminately; he discriminates among them, giving them different judgments to mark the differences between them. The bad men he punishes (albeit less than the rules prescribe), the reformable man he reforms, and the good man he lets go.⁸⁵ The ruler is merciful to all these men, each in the proper way, by giving a just punishment that is less than the social rules prescribe. Mercy may appear indiscriminate or haphazard with respect to the social rules, but it is continuous with discriminating justice. Mercy, which is continuous with justice, is the aim in responding to *all* crimes in a society that has sorely corrupt social rules about punishment.

Given this basic explanation of Seneca's concept of mercy, I turn to highlight the features that distinguish it from negative mercy. Unlike negative mercy, Seneca's mercy is governed by reasons, it is not motivated by pity or compassion, and it is continuous with justice. Although the language of moral requirements and supererogation are not in Seneca's stoic vocabulary, it is clear that he thinks mercy is not optional; insofar as a ruler is fully virtuous, she must be merciful toward all. I have mentioned some of these features in passing above, but I consider them in turn here in order to highlight the contrast with negative mercy.

⁸⁵ "To one man he will give simply a verbal caution, sparing him from any punishment if he is young and capable of reforming; another who is evidently labouring under the shame of his crime will be told to go free, because he was led astray, because wine caused him to fall," ("On Mercy," in *Dialogues and Essays* trans. Davie, 217).

Unlike most contemporary authors writing on negative mercy, Seneca maintains that mercy is reason-governed; “mercy operates together with reason.”⁸⁶ As discussed above, merciful judgment is discriminating and draws distinctions between different offenders; “the mercy we exercise ought not to be indiscriminate.”⁸⁷ When Seneca recommends mercy in a given case, it is for some reason. Also unlike contemporary writers about mercy (and unlike his predecessor Cicero), Seneca holds that mercy is not—indeed must not—be motivated by pity, or any other emotion like compassion or sympathy. Mercy is proper mental control in exacting punishment; pity is “the defect of a small mind... it is mental sorrow caused by the sight of others’ wretchedness...but no sorrow falls on the wise man; his mind is tranquil, and nothing can happen to cast a cloud over it.”⁸⁸ When a person feels pity, she cannot properly control and exercise her mind. Given that mercy is a kind of proper mental control, a person who is feeling pity cannot grant mercy.⁸⁹

⁸⁶ Ibid., 215.

⁸⁷ Ibid., 190.

⁸⁸ Ibid., 216.

⁸⁹ There is one sentence that seems to contradict my interpretation: “One man’s youthful years moves me, another’s advanced age; one I have pardoned for his high standing, another for his low; whenever I discovered no excuse for pity, I spared the man for my own sake,” (“On Mercy,” in *Dialogues and Essays* trans. Davie, 189). This passage is problematic in three respects. First, the passage conflates mercy with pardon; a distinction that Seneca insists on later in his letter. Second, it suggests that mercy is motivated by pity; a claim that Seneca flatly rejects later in his letter. Third, the end of the passage suggests that a ruler can grant mercy for no reason; the ruler can simply grant mercy because he feels like doing so, for his “own sake.” This sentence is clearly in tension with my interpretation. The third problem is easiest to address. Seneca writes that the ruler is “the soul of the state and state is [his] body,” (ibid., 192). Bearing this in mind, we can read the problematic remark as follows, ‘I spared the man for the sake of *the state*.’ This makes sense because Seneca thinks that the well being of the state is a

In contrast to contemporary writers about mercy, Seneca makes no suggestion that mercy is a departure from justice. Instead, Seneca maintains that mercy is continuous with justice;⁹⁰ mercy acts as though “the greatest justice resided in that action it has decided upon.”⁹¹

Although Seneca does not use the language of moral obligation or requirement, it makes sense for us to infer from his view that people, insofar as they are fully virtuous, must respond to offenders with mercy.⁹² Mercy is a must, not an option.⁹³ In a small

perfectly good reason to show mercy. The other two problems are not as easily resolved, but I try to address them as follows. The sentence appears at the beginning of “On Mercy,” well before Seneca draws firm distinctions among mercy, pity, and pardon. Given the sentence’s early position in the work and the firm distinctions drawn later, Seneca can be charitably interpreted as speaking loosely in this sentence. Seneca indeed suggests that people use terms like “clementia” loosely and tend to run it together with different but related concepts like pity and pardon. Seneca prefaces his definition of mercy by remarking on this looseness in language; “to avoid being sometimes deceived by the attractive name of mercy and directed into an opposite quality, let us see what mercy is, what its nature is, and what its limitations,” (ibid., 214). He makes a similar remark elsewhere when explaining why he bothers to define mercy; “there are certain vices that simulate virtues [and] they cannot be distinguished unless you stamp them with marks that enable you to tell them apart,” (ibid., 190). Early in the work, before Seneca has used his “stamp” to mark off mercy from pity and pardon, it is understandable for him to use these terms loosely.

⁹⁰ Several of Seneca’s contemporaries and immediate predecessors use “justitia” and “clementia” as synonyms to mean righteous behavior or behavior that is in accord with divine law. See Julius Caesar, *Commentaries on the Gallic Wars* (circa 50 BC). Also, in *History of Rome* (circa 25 BC), Livy combines justitia and clementia and refers to “clementiae justitiaeque.” Although Seneca and several of his contemporaries sometimes conflate “justitia” and “clementia,” they generally do not conflate these terms with “jus.” “Jus” refers to behavior that is obligatory (that one is bound to do, that one must do, that others can rightly force one to do)—whereas justitia and clementia refer to behavior that is righteous (that one should do, but which others may not rightly force one to do). See *Oxford Latin Dictionary*. I am indebted to Jim and Eleanor Leshner for help on this point.

⁹¹ “On Mercy,” in *Dialogues and Essays* trans. Davie, 217.

⁹² The moral requirement to show mercy does not, however, entail a moral right to mercy. An offender cannot demand that the sovereign grant mercy as a matter of right because

extract from the lost third book of “On Mercy,” Seneca boldly writes, “he who does not remit the punishment of wrong-doing is a wrong-doer.”⁹⁴ Mercy is a necessary virtue in a “blameless soul” and it is even a moral must for Caesar.⁹⁵ Although no other people could hold Caesar⁹⁶ accountable for granting mercy, Seneca argues that the gods would do so and his conscience would do so (as long as it is functioning properly). “On Mercy” opens with an “inspection of a good conscience” of a ruler who has kept “mercy always

the offender cannot demand anything of the sovereign by right. The sovereign is his superior; the sovereign has the power to decide his fate. Seneca does not defend a right to mercy because he does not defend any rights against the sovereign. He may have grounds for objecting to the general idea of rights against a sovereign, but he has no specific grounds for objecting to a right to mercy.

⁹³ One might think that a passage at the end of the text challenges this point, but the passage actually confirms it when we interpret freedom properly. Seneca writes: “Mercy has *freedom to decide*; not the letter of the law, but what is fair and good determines the sentence it passes,” (“On Mercy,” in *Dialogues and Essays* trans. Davie, 217). The freedom here is not freedom in the sense of having multiple options that are all morally right. The freedom here is freedom in the Stoic sense of judging on the basis of right reason. The merciful judgment follows directly from right reason; it is *determined by* what is fair and good.

⁹⁴ Extract preserved in a letter by Hildebert of Tours (“On Mercy,” in *Moral Essays*, trans. Basore, 449). The quote may be somewhat misleading because it may suggest that mercy is giving less punishment than is justly due. In context, however, it seems that mercy is refraining from giving an overly harsh punishment that one has the social power to give and giving instead only as much punishment as needed. “It is the part of mercy to cause some abatement of a sentence that aims at revenge. He who does not remit the punishment of wrong-doing is a wrong-doer. It is a fault to punish a fault in full. He shows himself merciless whose might is his delight. It is a shining virtue of a prince to punish less than he might. It is a virtue to be forced by necessity to take vengeance, not to visit it voluntarily,” (ibid.).

⁹⁵ “On Mercy,” in *Dialogues and Essays* trans. Davie, 189.

⁹⁶ Caesar’s social power was supreme; everyone was his social inferior and no one could hold him accountable for his decisions. He decided and acted with impunity from others. For discussion of how theories of kingship influenced Seneca’s work, see Braund, *De Clementia*.

in readiness.”⁹⁷ Only the merciful ruler is “prepared to give the immortal gods a full tally of the human race, should they require a reckoning.”⁹⁸ Seneca is not concerned about the existence of moral constraints on Caesar. He embraces the constraints as part of godly power rather than arguing that they improperly restrict sovereign power.⁹⁹

On Seneca’s account of mercy, mercy is a virtue consisting in the inclination to punish justly rather than in accord with social rules that call for harsher treatment. Drawing heavily from this account, I turn to spell out the concept of positive mercy.

Section III: Positive Mercy

Using Seneca’s conception of mercy as a starting point, I develop an account of positive mercy in this section. I describe the ways in which my account departs from Seneca’s view of mercy. Then I compare positive mercy to related concepts like equity, justice, and negative mercy.

Positive mercy is treating a person justly when unjust social rules call for harsher treatment (generally in the form of punishment). Positive mercy is, like Seneca’s mercy, continuous with justice, governed by reason, morally required, and not necessarily motivated by compassion or other forms of special concern.

Positive mercy departs from Seneca’s account of mercy in five respects. First, I define positive mercy as an action-type rather than a virtue. I take this departure not on principled grounds, but because doing so is apt for my overall project. Second, although

⁹⁷ “On Mercy,” in *Dialogues and Essays* trans. Davie, 188.

⁹⁸ *Ibid.*, 189.

⁹⁹ In such constraints “lies the slavery inherent in supreme greatness: it is unable to become less great; but you share with the gods this inescapable position,” (“On Mercy,” in *Dialogues and Essays* trans. Davie, 195).

positive mercy is generally exercised with respect to punishment, it is not necessarily limited to the punitive context. For example, a knight can show positive mercy upon his opponent at the end of sword fight. Given that my focus is on the punitive context, however, this is a very minor point. Third, the exercise of positive mercy is not incompatible with (but also does not require) feelings of compassion, sympathy, or pity. I neither want nor need to defend Seneca's stoic rejection of pathos in order to defend positive mercy. Fourth, Seneca maintains that mercy involves treating a person justly, but he sometimes takes a very broad, amorphous view of justice. He sometimes folds into justice considerations that we would not take as relevant to justice; for example, *raisons d'état* and maintenance of family loyalty. When I say, however, that positive mercy is treating a person justly, I mean to appeal to the more richly textured modern concept of substantive justice.

I tweak Seneca's account in a fifth respect. According to my view, exercising positive mercy is treating a person justly when *unjust* social rules (rather than *any* social rules, just or unjust) call for harsher treatment. For the purpose of the dissertation, I define an unjust social rule as a rule that is incompatible with a full theory of political justice. That is, the full theory rejects the rule based on a principle of justice. I contrast unjust social rules to decent rules. A decent social rule is compatible with (that is not ruled out by) a full theory of political justice.¹⁰⁰

The exercise of positive mercy treats a person with justice in derogation from an *unjust* social rule that calls for harsher treatment. I depart from Seneca's account in this

¹⁰⁰ I generally refrain from using the term "just rule." This label connotes (to me at least) that the rule is entailed by—rather than simply compatible with—a theory of political justice.

respect in order to distinguish positive mercy from the closely related concept of equity that Aristotle defends and contemporary literature and practice draws upon.¹⁰¹ On Seneca's account, mercy and equity are easily run together. On my adapted account, the distinction remains securely in place and can help provide a more richly textured understanding of positive mercy and equity.

Equity, argued Aristotle, is tailoring judgments to the particular features of a case rather than simply rendering judgments based on a strict application of rules.¹⁰² In his defense of equity, Aristotle maintains that justice depends on particularities of cases that laws, given their generality, cannot capture. The generality of rules is not an imperfection or failing of the rules themselves that can somehow be fixed;¹⁰³ rules must be general in order to be rules. Due to their generality, however, rules do not perfectly track justice and so derogation from the rules is sometimes required to achieve justice. Equity is derogation occasioned by a gap between the generality of rules and the particularities to which justice is sensitive.

The exercise of both equity and positive mercy involves departing from rules in order to act in a substantively just manner. Equity and positive mercy are different, however, in (at least) two respects. First, equity has a broader scope or extension than positive mercy. The exercise of positive mercy is limited to the modulation of harsh

¹⁰¹ Nussbaum, "Equity and Mercy."

¹⁰² *Nicomachean Ethics*, trans. Irwin, Bk. 5, Ch 10, 1137b 10-35. Aristotle limits his discussion of equity to laws. His discussion, however, can be applied to all (other) social rules.

¹⁰³ As Aristotle puts it, "the source of the error is not the law or the legislator, but the nature of the object itself, since that is what the subject matter of actions is bound to be like," (ibid., 1137b 15).

treatment (generally punishment) in the downward direction. The exercise of equity can give a person less *or more* harsh treatment than rules prescribe in a given case. A judge can also exercise equity with respect to decisions that do not necessarily involve the modulation of punishment. A judge can exercise equity whenever she is in a position to uphold any kind of rules, which due to their generality, call for an unjust decision. For example, a judge can exercise equity in making decisions about the admissibility of evidence and various points of trial procedure. A judge generally cannot exercise positive mercy when making these kinds of decisions because they generally do not involve modulation of punishment.

The more important difference for my purpose, however, is that the exercise of equity derogates from rules on different grounds than the exercise of positive mercy. An exercise of equity locates the grounds for derogation in the complexity of the specific case to which the rules are being applied whereas an exercise of mercy locates the reason for derogation within the rules themselves.¹⁰⁴ Equitable derogation from rules is occasioned by the gap between justice and the necessary generality of rules; merciful derogation, on the other hand, is occasioned by the gap between justice and the unnecessary imperfections in rules. In Aristotle's words, equity is a "rectification of law insofar as the *universality* of law makes it deficient" (emphasis mine).¹⁰⁵ But positive mercy is a rectification of law insofar as the *cruelty or injustice* of law makes it

¹⁰⁴ I am indebted to Emily Kelahan for her help in formulating this distinction.

¹⁰⁵ *Nicomachean Ethics*, trans. Irwin, Bk. 5, Ch 10, 1137b 25.

deficient.¹⁰⁶ The general/particular problem to which equity responds is a necessary feature of rules, but the injustice to which positive mercy responds is not necessary. Unlike positive mercy, equity would not be an endangered species in a society with perfectly just rules. Equity, guaranteed its place in the best possible polity, is on the mind of Aristotle the philosopher. Mercy, mired in the muck of imperfection, is drawn from the stuff of Seneca the jurist and politician.¹⁰⁷

In a criminal justice system with perfectly just rules (or even all and only decent rules), positive mercy would be obsolete. Like an appendix in the human body, it would be a non-functional vestige of a less perfect age. Beccaria seems to agree, maintaining that mercy should be excluded if the penal code were made perfect.¹⁰⁸ Interestingly, only

¹⁰⁶ Jeremy Bentham recognizes a similar distinction, but puts it in terms of interpreting the will of the legislator and he does not use the labels equity and positive mercy.

Interpretation may be distinguished into strict and liberal. It may be stiled strict where you attribute to the legislator the will which, at the time of making the law, as you suppose, he really entertained. It may be stiled liberal where the will you attribute to him is not that which, you suppose, he really entertained, but a will which, as you suppose, he failed of entertaining, only through inadvertency: insomuch that, had the individual case which calls for interpretation been present to his view, he would have entertained that will, which, by the interpretation put upon his law, you act up to, as if it had been his in reality. [*Equity involves this kind of departure from the letter, but not the spirit, of the law.*] I say through inadvertency: for to attribute to the legislator a will which you suppose him to have failed of entertaining though any other cause than inadvertency, that is from wrong judgment or perverse affections, and to act accordingly, is not to interpret the law, but to act against it. [*Positive mercy involves this kind of departure from both letter and spirit of the law.*] (*Limits of the Penal Branch of Jurisprudence*, Section 16, "Idea of a Compleat Law," 130)

¹⁰⁷ Seneca received mercy from Caligula who had intended to execute him on grounds of having aroused his jealousy. Later, Seneca was not so lucky. Ironically, he was wrongly sentenced to death by Nero, the man to whom he had addressed "On Mercy." He was forced to commit suicide and chose to do it by reenacting the last hours of Socrates. See Reinhardt, "Introduction," in *Dialogues and Essays* trans. Davie.

¹⁰⁸ Becarria writes:

one regime has ever officially abolished mercy—and that regime was radically optimistic about the prospect of achieving perfect justice. The unique and short-lived experiment occurred in 1789 during the French Revolution.¹⁰⁹

Before moving on to compare positive mercy to negative mercy, I would like to briefly describe the relationship between positive mercy and justice (I say more about this relationship in the next chapter). Following Seneca, I maintain that positive mercy is a continuous part of justice. It is not, however, equivalent to the whole of justice. All exercises of positive mercy are exercises of justice, but not all exercises of justice are necessarily exercises of positive mercy. Justice involves a great deal more than correctly doling out (or withholding) punishment and other harsh treatments. It involves, for example, developing and maintaining institutions that respect civil rights, treating people as free and equal persons, and fairly distributing goods and services.

As punishments become more mild, clemency and pardon are less necessary. Happy the nation in which they will be considered as dangerous! Clemency, which has often been deemed a sufficient substitute for every other virtue in sovereigns, should be excluded in a perfect legislation, where punishments are mild, and the proceedings in criminal cases regular and expeditious. This truth will seem cruel to those who live *in countries, where, from the absurdity of the laws, and the severity of punishments, pardons, and the clemency of the prince are necessary.* (“On Crimes and Punishments,” 97-98; emphasis mine)

A similar view is endorsed by Filangieri: “If a pardon is just the law is wrong, but if the law is not wrong a pardon is,” (as quoted in Walker, “The Quiddity of Mercy”). For modern defenders of this general kind of view, see Leslie Sebba, “The Pardoning Power;” Walker, “The Quiddity of Mercy,” and Pearn, “The Quiddity of Mercy—A Response.”

¹⁰⁹ Moore, *Pardons*. Moore puts the point in terms of “pardon” rather than “mercy,” but given my definition of positive mercy and her definition of pardon, the two are roughly the same. Moore explains that pardoning was abolished in 1789, but ad hoc procedures for clemency developed over the course of the following ten years. The pardoning power was formally granted to First Consul in 1802.

Within the more limited context of justice in criminal punishment, positive mercy might be extensionally equivalent to justice, but (I hope) this is a rare state of affairs. One can only exercise positive mercy in a case if unjust social rules (these include but are not limited to laws) call for overly harsh punishment in that case. One can exercise justice regardless of whether the relevant social rules are unjust or not. If all the social rules bearing on criminal punishment are unjust in a society (as perhaps they were in Seneca's time), then positive mercy would be extensionally equivalent to justice within the context of criminal punishment. I hope such a state of affairs is the exception and not the rule in modern political communities!

With the account of positive mercy in place, I turn to compare it to negative mercy. These two concepts of mercy do not have much in common. Negative mercy is sparing from deserved harsh treatment and positive mercy is derogating from rules to give just treatment (generally in the form of punishment). Negative mercy is distinct from justice and positive mercy is continuous with justice. Negative mercy is supererogatory, not necessarily justified by reasons, and driven by emotions like pity. In contrast, positive mercy is morally required, justified by reasons, and not necessarily driven by emotions like pity. Positive mercy cannot be exercised unless *unjust social rules* call for excessively harsh treatment toward some people. Negative mercy can be exercised regardless of whether operative social rules call for too little, too much, or just enough harsh treatment toward offenders. The concept of negative mercy is parasitic on the concept of desert (usually "just desert"), but positive mercy has potential to float free from desert (insofar as just treatment can be determined without an appeal to desert). Given that negative mercy is a departure from justice and is nevertheless good, it

presupposes that *justice* sometimes calls for excessively harsh treatment toward offenders. Positive mercy presupposes no such deficiency in justice. Positive mercy departs from social rules and presupposes that social rules are somehow deficient, but negative mercy presupposes that justice is somehow deficient.

Given the stark contrast between positive and negative mercy, one might object to my labeling of positive mercy as *mercy*. The concept is so different from the familiar concept of negative mercy that one might not want to call it “mercy” at all.¹¹⁰ I defend the label “positive mercy” on the basis of three reasons. First, historical evidence suggests that many people (other than Seneca) have used the term “mercy” to refer to the concept of positive mercy or something closely akin to it.¹¹¹ Second, there are many actions that we call “mercy” today that are instances of positive, rather than negative mercy. For example, people commonly say that pirates, black knights, and even Nazis show (positive) mercy when they spare their victims from harsh treatment.¹¹² Surely this

¹¹⁰ There are really two worries with the label “mercy.” The first worry, which I address here, is “why does positive mercy share the term ‘mercy’ with negative mercy?” The second worry is “why not call an exercise of positive mercy an exercise of plain old fashioned justice?” I address the second worry in the Conclusion of the dissertation.

¹¹¹ “The prerogative of mercy in the English legal system, for example, was used for centuries to distinguish the punitive treatment of those who had killed intentionally from those who had done so accidentally, a distinction not drawn by the law of murder,” (Tasioulas, “Mercy,” 113).

¹¹² A BBC News article reports the following incident as a case of mercy (see “Glass of Milk ‘Saved’ Nazi Victim”). A visibly pregnant Jewish woman was on a train headed to a concentration camp during World War II. En route to the camp, the train stopped in a small farming village and everyone was forced to get off the train. The pregnant woman, who had not eaten anything in three days, looked extremely ill. A farmer who happened to be in the area reached out to give the woman a glass of milk. A Nazi guard saw the farmer reaching out to the woman. According to the rules, no passengers were allowed to receive food or beverages so the guard lifted his arm as if to strike the woman or prevent the farmer from giving her the milk. But the guard did nothing. He stood by and

is not negative mercy; their victims do not *deserve* the harsh treatment from which they are spared.¹¹³ Third, negative and positive mercy share two important similarities despite their differences. Given these similarities, it makes sense for them to share the term “mercy.” First, both negative and positive mercy involve a powerful person withholding harsh treatment (generally punishment). Second, negative and positive mercy are both thought to flow from a recognition of human imperfection; negative mercy flows from a recognition of the frailty of human individuals (no one is perfect, we easily stumble, “there but for the grace of God go I”) and positive mercy flows from a recognition of the frailty of social rules and human institutions (social rules are rarely perfect, most institutions fail to be perfectly just).¹¹⁴

These reasons support my decision to call the concept that I have adapted from Seneca’s work “positive mercy.” If the reader remains unconvinced by my three reasons, I invite her to substitute the term “positive mercy” with “clemency” (from Seneca’s “*clementia*”). I do not want the label to impede the argument of the next section and chapter: the use of negative mercy faces several challenges in the context of criminal

watched the woman gratefully receive and drink the milk. He never struck or said a word to the woman or the farmer. The article calls both the Nazi’s and the farmer’s actions “mercy.” Both are positive, not negative mercy.

¹¹³ Rainbolt makes this point in “Mercy: An Independent, Imperfect Virtue.”

¹¹⁴ The etymology of the term “mercy” is perhaps a fourth reason to call both negative and positive mercy “mercy.” The etymology of “mercy” is often traced to the Latin “*miser cordia*” which means pity or sorrowful heart at the sight of another’s suffering. This etymological root supports the use of the term “mercy” in the label “negative mercy.” There is another root that supports the use of the “mercy” in the label “positive mercy.” The etymology of “mercy” is also traced (by way of the French “*merci*”) to the Latin “*merces*” which means giving what is owed, payment that is due, or treatment that is deserved. *Oxford Latin Dictionary*.

justice systems, but what I call positive mercy is apt to play an important role in those systems.

Section IV: Stumbling Blocks for Negative Mercy

Having identified and contrasted negative and positive mercy, I turn to compare their respective merit in the context of a criminal justice system.¹¹⁵ In this section, I present two problems for using negative mercy in a criminal justice system. These problems are not meant to blow negative mercy out of the water in criminal justice, but to raise important challenges to it. Instead of (or while we are) holding our breath for a resolution to these challenges, I recommend that criminal justice systems investigate and potentially revive positive mercy. Positive mercy is not susceptible to the problems of negative mercy and, as I discuss in the next chapter, has a great deal of merit in its own right.

The first problem with the exercise of negative mercy in a criminal justice system has its roots in a longstanding concern about the independence of mercy from justice.¹¹⁶ Negative mercy is defined as being independent from justice. But if mercy is indeed

¹¹⁵ By “criminal justice system,” I do not mean any centralized authority that engages in judging and punishing a set of offenses. I mean a public institution that is charged with achieving justice in response to a set of offenses. I assume that a public institution charged with achieving justice is obliged to treat like cases alike. Any criminal justice system, in the way I am using the term, is committed (at the very least) to the broad aim of justice and to the more particular aim of maintaining formal equality (that is, treating like cases alike).

¹¹⁶ This problem has its roots in a puzzle that was first articulated by St. Anselm. The puzzle has been dubbed “the paradox of mercy” and it has been the focus of much concern in philosophical literature. See Anselm, Murphy, Rainbolt, Tasioulas, and Brien. Anselm resolves this puzzle by maintaining that mercy is part of justice. Indeed mercy is the very root of justice according to Anselm. In this respect, Anselm appears to be working with the concept of positive mercy rather than negative mercy. I am indebted to Marilyn Adams for insightful conversation on this point.

independent from justice, then mercy requires some departure from justice. A departure from justice, however, is unjust. We should not act unjustly, so we should not show mercy. A common response to this problem is to suggest that not all departures from justice are unjust (they are perhaps a-just). This response may be helpful in some contexts, but it is unsatisfying in the context of a criminal justice system. Doing justice is the criminal justice system's *raison d'être*. A system that exists to do justice should not depart from justice, regardless of whether a departure from justice is strictly speaking unjust or simply a-just. We should not advocate for departures from justice in the very pursuit of justice. Let the departures happen elsewhere.

The second problem with negative mercy in criminal justice arises from a conflict between the supererogatory nature of negative mercy and the requirement of formal equality in criminal justice. I assume, not implausibly, that officials acting in the criminal justice system are obligated to treat like cases alike. Suppose that two defendants, D1 and D2, have committed similarly serious crimes, causing similar harms, with similar kinds or levels of culpability. In order to rightfully show mercy to D1 and not D2, a relevant reason must distinguish D1 from D2. Otherwise mercy would violate the obligation to treat like cases alike. Suppose a good reason does distinguish D1 from D2 and a judge grants mercy to D1. Now suppose the judge encounters D3 who has committed a crime with similar seriousness, harm, and culpability as D1 and D2. The reason that distinguished and grounded mercy for D1 also applies to D3 (and D3 has no other features that relevantly distinguish him from D2). On pain of failing to treat like cases alike, the judge is obligated to grant mercy to D3. But negative mercy is supererogatory, not obligatory; it is never morally required. Unless criminal justice

systems relinquish the obligation to treat like cases alike, they cannot rightfully accommodate the exercise of negative mercy.¹¹⁷

The features of negative mercy generate these two problems. Positive mercy is immune from these problems because the concept of positive mercy lacks the tension-generating features of negative mercy. The second problem does not arise because positive mercy is reason-governed and morally required rather than morally optional. The first problem does not arise because positive mercy is part of justice. Positive mercy has moral value not because it tempers justice, but because it is part of justice. Positive mercy's particular aim, which is part of the broader aim of justice, is to correctly temper unjust social rules and give offenders just punishment.

My point, however, is not to show that positive mercy outperforms negative mercy in problem contests. Contemporary philosophers have crafted interesting responses to these (and other) problems with negative mercy in criminal justice. The responses generally fall into one of two camps. Both camps agree that the use of negative mercy—as understood on the received view—poses initial problems and challenges in the context of criminal justice. The optimistic camp maintains that the concept of negative mercy can be adequately fixed for use in the criminal justice system by altering or abandoning one (or more) of the features of the received view.¹¹⁸ The

¹¹⁷ Duff, “The Intrusion of Mercy,” and Harrison, “The Equality of Mercy” discuss versions of this problem. Anselm also worries about this kind of problem. Although he finds a resolution to the first problem, he asserts that we cannot resolve the second problem: “*No reasoning can comprehend why, from those who are alike in wickedness, you [God] save some rather than others through your supreme goodness and condemn some rather than others through your supreme justice,*” (*Proslogion*, Ch. 11, 14).

¹¹⁸ The following authors argue for rejecting (or reinterpreting) one of the features of the received view of mercy. Johnson and Card reject the idea that mercy is distinct from

pessimistic camp retorts that the relevant feature(s) of mercy cannot be abandoned without changing the topic or improperly distorting the concept of mercy.¹¹⁹ Unwilling to discard any of the features of mercy, the pessimists generally respond to the puzzles by recommending that we resign mercy in the criminal court¹²⁰ to the trash heap of archaic political ideas.

Rather than adjudicating between these camps and declaring a victor, I propose a compromise between them. I maintain that the optimists are on to something when they defend an altered kind of mercy in the criminal justice system. Instead of defending what the pessimists take to be a contorted version of negative mercy, I suggest (re)adopting positive mercy for use in the contemporary criminal justice system. It is a different concept of mercy, but it has a venerable lineage and, as I will show in the next chapter, it is still a useful concept today. I am offering an olive branch in the form of an alternative positive view, not indicting either side.

The compromise I propose neither proves or denies that negative mercy is apt for use in the context of criminal justice. Some may want to further analyze negative mercy and make room for it in criminal justice in spite of (or in light of) its problematic conceptual features. I have no “knock-down” objection to such an approach, but it is not

justice. Smart, Walker, Brien and Tasioulas reject the supererogatory nature of mercy, holding that mercy is morally required when certain reasons apply. Walker and Tasioulas also reject the view that mercy is not reason-governed. Still other authors are inclined to make distinctions within kinds of mercy, holding that mercy is sometimes obligatory and sometimes supererogatory (see Sterba and Rainbolt).

¹¹⁹ See Hurd, Markel, Statman, Harrison, Nuyen, and Murphy, *Forgiveness and Mercy*.

¹²⁰ Most critics refrain from making the strong claim that mercy *tout court* should be resigned to the trash heap. They generally do not object to the use of mercy in the context of interpersonal interaction (Hurd), private law (Murphy), and theology (Nuyen).

my own. My argument is that discussion of mercy in the context of criminal justice should not be limited to negative mercy. Positive mercy not only lacks the problematic conceptual features of negative mercy, but as I argue in the next chapter, it also has a great deal to say for itself in its own right. I will not show that it is wrong to stop investigating the use of negative mercy in criminal justice, but that it is myopic to carry on as if negative mercy were “the only game in town.”

CHAPTER 2: DEFENSE OF POSITIVE MERCY

“The American Bar Association should consider a recommendation to reinvigorate the pardon process at the state and federal levels. The pardon process, of late, seems to have been drained of its moral force. Pardons have become infrequent. A people confident in its laws and institutions should not be ashamed of mercy.”

~US Supreme Court Justice, Anthony Kennedy¹²¹

Although modern criminal justice systems respond to crime in a considerably more humane fashion than the Roman Empire, they still harbor some unjust social rules that call for excessively harsh punishments. The existence of such unjust rules give rise to ample *occasions* for the exercise of positive mercy, but it is unclear whether officials in criminal justice systems should actually exercise positive mercy when such occasions arise. If, for example, an unjust law requires a judge to give an offender an excessively harsh sentence, should the judge derogate from the law and give the offender a just sentence?

Seneca does not seem to think twice about advocating for the use of mercy. After all, exercising positive mercy gives substantively just treatment to offenders in individual cases. A modern audience, however, should think twice about directly importing Seneca’s approach into the modern context. Seneca lived in a context in which formal justice under law was virtually non-existent; he was writing to Nero who was ruled by no one and who had absolute power. Officials in modern criminal justice systems (judges,

¹²¹ Address at the American Bar Association Annual Meeting.

lawyers, police officers, wardens, etc.) are in a very different position. They are *ruled by the laws* of the institutions in which they operate and through which they are given restricted power. Unlike a Roman Emperor, their proper aim is not to single-handedly mete out substantive justice in each case that comes before them. They are constrained by formal justice under law; their proper aim is generally understood to be achieving *justice under the rule of law*. The fact that positive mercy does substantive justice in individual cases does not necessarily give them, unlike Nero, an appropriate reason to show positive mercy if doing so stands in violation of law.

Seneca's work is of interest to a modern audience because (among other things) it directs our attention to positive mercy, but not because it articulates good reasons for modern government officials to exercise positive mercy. We fortunately live in a very different political climate than Seneca and his audience. If we are interested in using positive mercy in modern criminal justice systems, we must do our own work to defend it.

In this chapter, I aim to do this work. My objective is not to show that specific criminal justice systems should incorporate a practice of positive mercy in any particular fashion. Rather, my objective is to show that officials within criminal justice systems have standing reason to exercise positive mercy and that most political communities have standing reason to incorporate a general practice of positive mercy into their criminal justice systems. To see what reason stands in favor of positive mercy, I suggest how a general practice of positive mercy might work and explore some of its advantages and potential challenges. In Section I, I present an outline of how a political community might incorporate a practice of positive mercy into its criminal justice system. With this

outline in place, I begin my substantive argument for why officials have reason to exercise positive mercy and why political communities have reason to incorporate a practice of positive mercy in their criminal justice systems.

To argue that officials have reason to exercise positive mercy, I distinguish between two types of positive mercy: a) law-abiding positive mercy that requires derogation from a social rule that is not a law and b) outlaw mercy that requires derogation from a law. In Section II, I show that officials have reason to exercise law-abiding positive mercy. I also argue that political communities have reason to incorporate a general practice of such law-abiding positive mercy. In Section III, I take up the significantly more challenging question of whether officials have reason to exercise positive mercy when doing so requires them to derogate from law. By engaging in a substantial analysis of the value of the rule of law in modern political communities, I argue that judges have reason to exercise such outlaw positive mercy. I show that exercising positive mercy is a way for judges to do justice not outside of the rule of law, but in service to the rule of law. On the basis of this argument, I conclude that political communities have reason to incorporate not only a general practice of law-abiding mercy, but also a judicial practice of outlaw mercy. In Section IV, I consider and respond to a powerful objection to my argument. To conclude the chapter, I say how my discussion and defense of positive mercy in Part 1 of the dissertation leads to Part 2 of the dissertation.

Section I: Outlining a Practice of Positive Mercy

A criminal justice system that has incorporated a practice of positive mercy is different from one in which officials occasionally exercise positive mercy as “one-off”

acts. This section provides a tentative description of how a criminal justice system might incorporate a practice of positive mercy. The description is not meant as a policy recommendation; it has neither the detail nor the requisite analysis to support it. Rather, the description is meant to fill-out the potential role that positive mercy can play in a criminal justice system so that we can better analyze its potential merit.

A criminal justice system that has incorporated a practice of positive mercy will have basic guidelines about which officials can exercise positive mercy, when they should do so, and how they should do so. It should also have a procedure designed to hold officials accountable for their decisions to grant positive mercy.

Who? A political community can incorporate positive mercy into its criminal justice system by giving any number of officials the authorization to grant positive mercy. Almost any type of official with control over criminal sentencing is an eligible candidate for granting positive mercy. Not all types of officials are equally good or advisable candidates for showing positive mercy, but I list them all because they seem reasonably eligible for the task. Judges are perhaps the most obvious candidates because they play the most visible role in deciding criminal sentences. Juries can also be authorized to show positive mercy when they play a role in sentencing (as they do in capital cases in the American system). Prosecutors can also be authorized to show positive mercy when they are deciding which charges to press, which sentences to recommend, and which bargains to negotiate. Probation and parole officers are also eligible candidates; they can suggest mercy in their sentencing recommendations and they can grant mercy in deciding parole and probation conditions and in deciding on whether to report offenders who violate terms of parole or probation. Prison wardens and

guards may also be authorized to show mercy to offenders by improving conditions within the prison to make sentences less harsh. Police officers generally are authorized to show positive mercy by deciding whether or not to arrest someone. In some criminal justice systems, many of the above officials already have authorized discretion that they can use to show positive mercy.¹²²

In many political communities, the executive branch of government is given broad clemency and pardon power that it can use to exercise positive mercy with respect to criminal punishment. The executive branch, however, does not operate as a part of the criminal justice system. It exerts an important, but *external* influence upon it. As such, the executive use of positive mercy falls slightly outside my aim of investigating the use of positive mercy *within* criminal justice systems. That said, however, I agree with Anthony Kennedy (quoted above) that the executive pardon process should be reinvigorated. The arguments of this chapter can lend indirect support to the claim that executives have reason to make use of their pardon power to exercise positive mercy.

When? A criminal justice system that has incorporated a practice of positive mercy should give officials rough guidelines on when to show mercy. For example, the guidelines could recommend showing mercy when two conditions are met. The first condition is that an applicable law(s) and/or other social rule(s) is unjust and prescribes overly harsh punishment. The law/rule might be entirely unjust (no amount of exceptions can make the rule just—for example, Aryans are superior to Jews) or they might be partially unjust (the rules can be salvaged if a few more exceptions are built into it). The second condition is that the official has it within her power not only to withhold an

¹²² Kadish and Kadish, *Discretion to Disobey*.

excessively harsh treatment from an offender, but also to respond to offender in a way that is just (or at least significantly closer to just).¹²³

How? A criminal justice system that has incorporated positive mercy should have rough guidelines on how officials should show mercy. These guidelines may widely vary from official to official (the way a judge ought to show mercy may be very different from how a prosecutor ought to show mercy). As will become clear for reasons that I discuss in the next two sections, I think the guidelines should recommend that officials show mercy publically; that is, in a fashion that is open and transparent to all members of the public. When possible, officials who grant mercy should publically articulate their reasons for granting mercy (and the reasons should be public reasons rather than the official's personal preferences). Officials should assert that the rule from which they are derogating is unjust and explain why it is unjust. They should also explain why the merciful response (whatever that might be) is just. I suspect that publicity will help get the most out of a practice of mercy and limit officials' abuse of the power to grant mercy.

Accountability? A criminal justice system that has incorporated a practice of positive mercy may have a procedure for reviewing officials' decisions to show mercy. A political community may hold officials accountable for mercy decisions by enacting a scheme that is specially tailored for periodically reviewing mercy decisions. Or it may

¹²³ This condition might often fail to obtain for members of a jury because the jury often has very limited options. Suppose a jury is only given two options to choose between: they can sentence a man to death or sentence him to only five years in a minimum-security prison. The jury decides that sentencing the man to death would be unjust, but they also decide that sentencing him for five years would be unjust (suppose he is a violent serial killer who is almost certain to strike again, and again). If sentencing the man to five years would be a worse miscarriage of justice than sentencing him to death, then the jury is not in an appropriate position to grant positive mercy. They should leave it to the judge who is empowered to give the man a just sentence.

hold officials accountable for mercy decisions within the general scheme(s) it already has in place for holding officials publically accountable. Either way, officials who reduce just sentences for bad reasons (like cronyism, nepotism, racism, or sexism) and glorify such miscarriages of justice with the name “mercy” should be held accountable. They may be held accountable for abusing their power in a variety of ways; their decision might be overturned (if that is an option), they may be publically criticized, denied promotions, fired, or even criminally punished. There is a great deal more to be said about how officials should be accountable for their decisions to show mercy, but I postpone this discussion until I have more carefully investigated the reasons that officials have to exercise mercy in modern criminal justice systems.

Section II: Law-Abiding Positive Mercy

The exercise of positive mercy in a criminal justice system grants a person a just punishment (that might mean *no* punishment, a shorter punishment, or an otherwise lighter punishment) when unjust social rules operative in the system would have called for harsher punishment. In order for a person to exercise positive mercy, she must derogate from some unjust social rule. Sometimes the exercise of positive mercy requires the granter of mercy to derogate from an unjust social rule that is not a law and sometimes it requires the granter to derogate from a law. In this section, I argue that officials have reason to exercise positive mercy when doing so does not require them to derogate from the law. In the next section, I analyze cases in which positive mercy does require officials to derogate from the law.

I begin by explaining when occasions for such law-abiding positive mercy arise. Sometimes criminal justice systems have unjust informal rules, policies, or non-legal

customs that call for overly harsh punishment toward criminal offenders. Non-legal rules exert a great deal of influence over the way that many criminal justice systems treat criminal offenders. For example, informal (or formal) social rules often govern the way in which prosecutors cut deals, how parole boards are expected to interpret and weigh various factors, how wardens run prisons, and the ways in which judicial discretion is appropriately used and abused. The law is often silent on whether officials should follow these rules or not. Suppose that one of the rules is unjust and that an official is not legally bound to follow it. If she deviates from the rule to give an offender just, rather than overly harsh treatment, the official exercises positive mercy in a law-abiding fashion.

Officials have two reasons to exercise positive mercy in such a law-abiding fashion. First, and most obviously, officials have reason to do justice under law and these exercises of positive mercy do justice under law. If an official were to withhold positive mercy when she had occasion to exercise it, she would deny an offender substantively just treatment when the law provided her with the opportunity to provide it. This kind of behavior might be acceptable in rare circumstances, but in general, officials should avoid it and show positive mercy when they have occasion to do so.

The achievement of substantive justice is also a reason for officials to exercise the closely related act of law-abiding equity. Sometimes a social rule demands excessively harsh treatment in a given case but the rule itself is not unjust. If the official is not legally bound to follow the rule and she derogates from it in order to give an offender just treatment, then the official exercises law-abiding equity. I do not, however, discuss equity in great depth here because my focus is on positive mercy.

In addition to doing justice under law, officials have a second reason to exercise positive mercy (this reason does not seem to apply to equity). The exercise of positive mercy is one way in which an official can fulfill her moral obligation to support just institutions. Following Rawls, I assume that members of a political community—officials as well as non-officials—have a pro tanto moral obligation to support and maintain existing, minimally just institutions within their community.¹²⁴ Part of what it means to support these institutions is to correct for injustice that is harbored within them. (Just as part of what it means to support and maintain one’s health is to root out disease harboring within it.) A person who conscientiously fulfills her duty to support just institutions keeps a critical eye out for injustices that might be lurking within (or glaring on the surface of) those institutions. Insofar as there is injustice within the institution, she stands in active opposition to it and acts in ways that are conducive to rooting it out. There are several courses of action she can take to do so. For example, she can lobby for, vote, or pass (if she is a legislator) new legislation; she can raise public awareness, or engage in symbolic protest.¹²⁵

Law-abiding positive mercy, I contend, is a way for officials within criminal justice systems to support the just institution¹²⁶ in which they work and which they generally know the most about. The act of positive mercy takes an active stand against

¹²⁴ The assumption is squarely shared by Rawls (*Theory of Justice*, Sec. 51-53). Rawls argues that individuals have a natural duty to support and comply with just institutions. Rawls thinks officials have a special duty (special in the sense that non-officials do not have it) to comply with law that is rooted in the duty of fair play.

¹²⁵ For insightful discussion of this method, see Hill, “Symbolic Protest and Calculated Silence.”

¹²⁶ I assume that most criminal justice systems are at least minimally just institutions.

unjust social rules in the criminal justice system, derogating from them because they are unjust and not worth having in the system. In many cases, the act of positive mercy not only takes a stand against an unjust rule, but also is conducive to calling forth positive change in the rule.¹²⁷ If positive mercy is strategically exercised in public or “on record,” it can add to (or establish) a body of practice (including but not limited to legal precedent) surrounding an unjust rule that can fund positive change of that rule. Like protests that are sometimes “only” symbolic but sometimes bring about significant reform, positive mercy will sometimes have no effect on changing the rules and sometimes will have considerable effect. Either way, the exercise of positive mercy is a way for an official to fulfill her obligation to support just institutions. This reason to exercise positive mercy stands in addition to the first reason that officials have to exercise law-abiding positive mercy: to do substantive justice under law.

Unfortunately, a variety of obstacles make it difficult for officials to recognize and act on occasions for positive mercy. Officials often internalize informal social rules and policies; they often take them for granted as correct and apply them without a second thought. Officials who do not exercise law-abiding positive mercy when they ought to do so might aptly be described as guilty of a callous (perhaps cruel) conventionalism or the maintenance of a brutal status quo. As history teaches us, it is easy for people to fall into this kind of behavior and often difficult for individuals to resist it. To exercise positive

¹²⁷ Pearn agrees that mercy is conducive to rooting out injustice in institutions and goes even further, arguing that “mercy is a necessary but not sufficient component for the evolution of a number of human institutions, of which the law is but one example,” (“Quiddity of Mercy,” 603). He takes rules of war as an example to illustrate his point. Acts of derogation from norms of brutality on the battlefield by powerful individuals developed into rules of chivalry in the middle ages and eventually into modern “jus in bello” codes.

mercy, officials must question the rules around them; doing so is often difficult. In addition to the psychological difficulty of questioning the rules, officials may also face social sanctions for questioning and derogating from social rules. Exercising positive mercy is a kind of “whistle-blowing” on bad practice in the criminal justice system and it comes with many of the general difficulties that whistle-blowers generally face.

Having a general practice of positive mercy would help officials engage in the difficult exercise of positive mercy. It would give a name to their practice and make it more visible for others to see (and be encouraged by). A general practice of positive mercy would also include an accountability scheme to protect individuals exercise positive mercy. Like laws and policies that are designed to protect corporate whistle-blowers, an accountability scheme for positive mercy would help protect those who have exercised positive mercy when they ought to have done so (the protection would prevent them from losing their jobs or facing other sanctions) and sanction those who abuse the practice by masking unjust decisions under the name of “mercy.”

Section III: Outlaw Positive Mercy

In this section, I argue that officials have reason to exercise positive mercy when doing so requires them to derogate from *law*¹²⁸ rather than (or in addition to) a non-legal

¹²⁸ By “laws” I mean laws in Dworkin’s post-interpretive sense. A law “on the books” may be a law in a pre-interpretive sense, but not in the post-interpretive sense. According to Dworkin, law is not equivalent to various rules listed “on the books”; rather, law requires interpretation. When a body of law has been interpreted in its best light, it is possible (indeed probable) that certain rules “on the books” cannot be understood as anything but mistakes or aberrations in the law. If what appeared to be a rule “on the books” does not fit in the soundest theory or interpretation of the law, then it is not a law in the strong, post-interpretive sense. It is only a law in the weak, pre-interpretive sense (it is a law only on its face, “prima facie,” but not a substantive law). Officials in most

social rule. Let me explain the circumstances under which occasions for such “outlaw” exercises of positive mercy arise. Many political communities have sentencing laws that not only allow, but require officials to give minimum punishments for certain offenses.¹²⁹ For my purposes, I call this kind of sentencing law unjust if it is incompatible with a full theory of justice. In virtue of its generality as a law, a just sentencing law may prescribe excessively harsh punishments in a haphazard handful of cases. Judges have occasion to exercise equity with respect to just laws, but they do not have occasion to exercise positive mercy. The exercise of outlaw positive mercy is a response to an unjust sentencing law.

Generally, an unjust sentencing law is incompatible with a full theory of justice because it has some defect in virtue of which it prescribes excessively harsh punishments in an identifiable pattern of (or perhaps all) cases.¹³⁰ By “excessively harsh” punishments, I mean punishments that involve more harsh treatment than is justifiable according to a fully theory of political justice. The defect could be, among other things, that the law fails to draw distinctions where distinctions ought to be drawn (for example, by prescribing the same punishment for mentally insane offenders as sane offenders who

criminal justice systems have a duty to uphold law in the strong, post-interpretative sense, but not in the weak, pre-interpretative sense. Dworkin, *Law’s Empire*, 107-108.

¹²⁹ All Western nations have passed legislation that removes or severely diminishes judicial discretion in criminal sentencing. Roberts, “Public Opinion and Mandatory Sentencing.”

¹³⁰ There may be other sentencing laws that are incompatible with a full theory of justice because they have some flaw that assigns too little punishment in a set of cases. These laws are not my focus here. My argument in the latter part of this section can be applied to show that derogation from these laws generally conflicts with the rule of law. In contrast, I argue that positive mercy serves the rule of law.

commit the same offenses), or that it draws distinctions where they should not be drawn (for example, by prescribing more punishment for black offenders than white offenders who commit the same offenses), or that it otherwise requires punishments that are morally unjustifiable (for example, by prescribing twenty-five to life sentences for petty theft like stealing a TV or some videos¹³¹). Without introducing any greater injustice, legislators could remove these laws or add exceptions to them to prevent them from prescribing overly harsh treatment toward types of offenders.¹³² Legislators should do so, but sometimes they do not. Through negligence, iniquity, political pressure, or other reasons, legislators sometimes let these unjust sentencing laws persist in the system.

These unjust laws are the focus of the remainder of this chapter. They leave officials in the criminal justice system in a sour predicament. We generally think that officials are charged to do substantive justice while remaining within the constraints of the law; they are to seek *justice under law*. The unjust laws described above, however, make it impossible for the official to give certain offenders a substantively just outcome while remaining within the constraints of the law. Should an official uphold these laws in cases where they require her to mete out excessively harsh punishment to offenders? Or should she derogate from them in such cases and deliver substantive justice to the offender who stands before her? If she takes the latter course of action, she shows

¹³¹ About 8,500 people in California are serving twenty-five to life sentences under the three-strikes law. About 3,700 of them are serving twenty-five to life sentences “for a third strike that was neither violent nor serious,” (Bazelon, “Arguing Three Strikes”). An army veteran with a heroin addiction is currently serving fifty years for stealing videos worth \$150 from a K-Mart. The Supreme Court affirmed his sentence in *Lockyer v. Andrade*, 538 U.S. 63 (2003). See also Pater, “Struck Out Looking.”

¹³² I am stipulating this fact in order to carve out the particular class of unjust laws that I am talking about in the rest of the chapter.

positive mercy. The question is: should she, or does she have any good reason to show positive mercy in such an outlaw fashion? (Henceforth in this section, I use “positive mercy” to refer to the exercise of positive mercy in such an outlaw fashion.)

In what follows, I argue that officials in the criminal justice system have standing reason to exercise positive mercy. First, I give what I take to be the best case against positive mercy: it seems to be at odds with the rule of law which has considerable instrumental and non-instrumental value in political communities with minimally just legal systems. Second, I argue that, contrary to appearance, positive mercy serves rather than threatens the rule of law in the criminal justice system. I conclude that insofar as officials have reason to pursue justice under the rule of law, they have reason to exercise positive mercy.

Rule of Law and the Case Against Positive Mercy

To begin, let me explain what the rule of law is and why it appears to conflict with positive mercy. Legal and political philosophers have developed slightly different conceptions of the rule of law. Lon Fuller maintains that the rule of law is an ideal that is achieved, to a greater or lesser degree, when the law satisfies eight principles: laws must be general, promulgated, prospective, clear, non-contradictory, possible to keep, constant, and most importantly for my purpose, *congruent with the behavior of officials of the regime*.¹³³ For Joseph Raz, the rule of law is the “principled, faithful application of the law,”¹³⁴ for Ronald Dworkin it is the “ideal of rule by an accurate public conception of

¹³³ Fuller, *The Morality of Law*. John Finnis has a very similar conception of the rule of law and endorses, for the most part, Fuller’s eight principles. Finnis, *Natural Law and Natural Rights*, 270-276.

¹³⁴ *Ethics in the Public Domain*, 373.

individual rights,”¹³⁵ and for John Rawls it is “regular and impartial, and in this sense fair, administration of law.”¹³⁶ Jeremy Waldron characterizes the rule of law very broadly as “law being in charge in a society.”¹³⁷ He maintains that the philosophical core of the debate over the rule of law is not really about what the rule of law *is*, but about how, if at all, to structure institutions such that law is in charge. Waldron’s approach has a great deal of merit and it guides me through the argument of this section.

Although legal and political philosophers disagree a great deal over how to structure an institution such that law is in charge, they do generally agree on at least one principle.¹³⁸ If law is to be in charge, political institutions must meet Fuller’s eighth principle. The rule of law requires government officials to uphold and apply the law as it is, not as they would like it to be. If the officials do not actually apply the law in practice, then the law is “a lame duck” or some other kind of powerless figurehead. Just as no number of crown jewels can make it the case that Queen Elizabeth II is in charge of modern Britain, so too no amount of generality, clarity, promulgation, or constancy can make it the case that law is in charge of a political community if its officials do not actually apply or enforce the law. Law does not have a magical scepter; it cannot rule unless there are officials charged to uphold and apply it. To use Fuller’s phrase, officials

¹³⁵ *A Matter of Principle*, 11-12.

¹³⁶ *Theory of Justice*, 38.

¹³⁷ “Rule of Law an Essentially Contested Concept,” 157.

¹³⁸ I have the following philosophers in mind: Fuller, Raz, Rawls, Finnis, Dworkin, Waldron (in most moods), and Colleen Murphy. Meir Dan-Cohen is an exception to the general trend. He argues that congruence is not necessary for law to rule and maintains that it is an open question as to whether there ought to be congruence between promulgated rules and official behavior (*Harmful Thoughts*, 38-40).

must have “fidelity to the law” if law is to rule. If officials toss aside laws that they do not like rather than uphold them, the laws are not in charge. The officials are in charge.

The case against positive mercy is beginning to take shape. An official who exercises positive mercy derogates from a law instead of applying it. Her exercise seems to stand at odds with Fuller’s eighth principle and the rule of law. But as it stands, this is not yet an argument against the exercise of positive mercy. So what if the exercise of positive mercy is at odds with the rule of law? To understand the best case against positive mercy, we need to see why the rule of law is worth having and maintaining in a given political community.

In what follows, I present four values of the rule of law in the context of a political community with a minimally just legal system.¹³⁹ I will return to analyze these values in much greater depth later in the argument, so I only briefly outline the four

¹³⁹ I limit the scope of my discussion to the value of rule of law in a minimally just legal system. In a wicked legal system, it is unclear whether the rule of law exists or has any instrumental or non-instrumental value. A great deal of the debate in literature on the rule of law bears on this interesting question (see, for example, Colleen Murphy, “Fuller and the Moral Value of the Rule of Law” and Dyzenhaus, *Recrafting the Rule of Law*). I do not, however, engage in investigation of this question because it is beyond the scope of my work here. There is no deep or general challenge to the exercise of positive mercy in the context of a wicked legal system. I assume that officials in wicked legal systems do not have a strong reason (perhaps no reason at all) to apply the wicked rules of that system. They do, however, have reason to try to achieve substantive justice for those who stand before them. Doing substantive justice is a sufficient reason for them (like Seneca’s audience) to exercise positive mercy. My focus is on minimally just legal systems where the rule of law is present and has considerable value. This is the context in which today’s liberal democracies are situated and in which positive mercy faces a deep challenge.

values here. Following the literature on the rule of law and its value, I group the rule of law values into two categories: instrumental values and non-instrumental values.¹⁴⁰

The rule of law has (at least) two instrumental values. First, the rule of law is instrumentally valuable because it is needed in order for law to do its job of guiding and coordinating human behavior. As Raz puts it, “the rule of law is essential for securing whatever purposes the law is designed to achieve.”¹⁴¹ Insofar as the laws in a minimally just legal system serve some worthwhile purposes (which they do in most modern political communities), the rule of law has instrumental value in helping the law achieve those purposes.

Second, the rule of law is instrumentally valuable because it keeps the law transparent to the public. Among other things, transparency afforded by the rule of law bolsters the public’s ability to reform the law (they cannot change a law if they do not know it exists) and it minimizes the chance (and the suspicion) that some people, usually the elite, have access to secret laws from which they can glean an unfair advantage.

The rule of law also has (at least) two non-instrumental values. First, the rule of law respects the autonomy of individuals in the political community. It respects each individual as a self-directed agent with the capacity to make decisions and plan her own life—as she sees fit, not as the ruling class deems fit. By ensuring that the ruling power is used according to stable, promulgated, and non-retroactive rules, the rule of law

¹⁴⁰ This categorization echoes Margaret Jane Radin’s approach in, “Reconsidering the Rule of Law.” She sees two categories of rule of law values: instrumental values and substantive values. I depart from her approach, however, by listing different values under each category.

¹⁴¹ *The Authority of Law*, 164

respects the capacity of an individual to guide her life by her own decisions. When the rule of law is in place, officials judge members of the political community according to “standards of behavior that they had a real opportunity to follow. Thus it is the decisions and actions of individuals—not the whims of officials—that determine the legal treatment that individuals receive.”¹⁴² The rule of law lays out a map upon which we can chart our own course; with the rule of law, we do not have to constantly guess which way to jump in order to avoid trouble from an arbitrary and unpredictable will. The rule of law is not just a tool that happens to bring about a state of affairs in which people can make decisions and plans based on reliable expectations (although that is good too!); in its very structure, the rule of law *respects our capacity* to make autonomous choices and plans. It treats us as self-directed, responsible agents.¹⁴³

Second, the rule of law structures a relation of reciprocity between those who rule and those who are ruled.¹⁴⁴ It is unfair for rulers to be unconstrained in their use of power while the ruled are kept constrained by that power. When law is in charge, those on both sides of political power are constrained. Waldron, picking up on Fuller’s recognition of the value of reciprocity in the rule of law, explains how this works.

¹⁴² Colleen Murphy, “Fuller and the Moral Value of the Rule of Law,” 250.

¹⁴³ Rawls argues that protection of liberty is the primary value of rule of law. He does not, however, use liberty in the sense of bare negative liberty (as Hayek tends to do in *The Road to Serfdom*). Rather, he uses protection of liberty in a sense that is compatible with what I call respect for autonomy. He writes, “the various liberties specify things that *we may choose to do*, if we wish, and in regard to which, when the nature of the liberty makes it appropriate, others have a duty not to interfere,” (*Theory of Justice*, 210; emphasis mine).

¹⁴⁴ I am indebted to Gerald Postema for encouraging my attention toward this value of the rule of law. I am also indebted to Colleen Murphy’s clear exposition and defense of this value in “Fuller and the Moral Value of the Rule of Law.”

Suppose the government wants to pursue goal G and decides to do it through rules (i.e. the rule of law). Officials are then bound to act according to the rules. They cannot act outside of the rules with respect to individuals even if doing so helps the government pursue goal G. Individuals then feel like they are being taken into account.¹⁴⁵ Non-officials are not the only ones constrained by law and forced to make sacrifices in light of them; officials who make and enforce the law also face such constraints and are forced to make sacrifices in light of them. The ruled and the ruling are *together* facing constraints of life in a political community. Under the rule of law, the ruled are not continually giving while the rulers continually take; instead, there is “give and take” on both sides.¹⁴⁶ As Fuller puts it, “there is a kind of reciprocity between government and the citizen with respect to the observance of rules. The government says to the citizen¹⁴⁷ in effect, ‘these are the rules *we expect you* to follow. If you follow them, *you have our* assurance that they are the rules that will be applied to your conduct” (emphasis mine).¹⁴⁸ This “kind of reciprocity” is certainly not tantamount to a full achievement of the political ideal of reciprocity, but it nevertheless instantiates an important aspect of that ideal.

¹⁴⁵ Waldron, “Why Law,” 278.

¹⁴⁶ Finnis also argues that reciprocity is a non-instrumental value of the rule of law. The rule of law has value in “holding rulers to their side of a relationship of reciprocity, in which the claims of authority are respected on condition that authority respects the claims of the common good (of which a fundamental component is respect for the equal rights of all to respectful conditions),” (*Natural Law and Natural Rights*, 274).

¹⁴⁷ Although Fuller uses the term citizen here, I have tried to avoid using it myself. Law’s rule holds over every resident of the political community, regardless of whether they have citizenship in the community. In a political climate where the law’s protection of non-citizens is under threat, it is particularly important to choose our words carefully so as not to encourage this threat.

¹⁴⁸ *The Morality of Law*, 39-40.

It is easy to miss the importance of the kind of reciprocity that the rule of law instantiates. In order to help us appreciate its importance, let us imaginatively enter into a regime where it was absent: Seneca's Rome. A myriad of political values and ideals were absent there and reciprocity is one of them. There was no fair give and take between ruler and ruled. If Seneca so much as ruffled the feathers of the Emperor, he would be executed. But if the Emperor committed any number of evil actions against Seneca, he would walk (or be carried) away with entire immunity. The emperor could do anything toward his subjects with impunity, but his subjects could do nothing without the threat of imperial sanction. "That's just *not fair*," we say (or exclaim) in our effort to articulate what is wrong with such a state of affairs. What is missing (among other things of course) is reciprocity in the relationship between the Emperor and the subject.¹⁴⁹ Waldron, Fuller, Finnis, and Murphy purport that the rule of law captures, in its very structure, the valuable reciprocity that I take to be missing here.

Equipped with an understanding of the rule of law and its value in, I can now make what I take to be the best general case against positive mercy. Modern political communities have created and strive to maintain the rule of law. Not without good reason, the rule of law has both instrumental value as a method of governance and it respects autonomy and reciprocity. Modern political communities give officials power on the condition that these officials use it in service of the rule of law. To maintain the rule of law in virtue of which they have been given their power, officials are obligated to faithfully uphold the law. They must follow Fuller's eighth principle and keep their

¹⁴⁹ The same reciprocity seems to be missing in Feinberg's Nowheresville. See "The Nature and Value of Rights."

behavior as officials congruent with the law. If a law is unjust, officials can use the legal channels that citizens use to seek reform. But when acting as officials, they are obliged not to derogate from law in some effort to “save the system from itself.” Officials are required to stay (and maintain) the path set by law, not off-road in their own Hummers of personal conviction. The exercise of positive mercy, however, seems to do just that. The official who derogates from law to show positive mercy seems to commit a kind of mutiny; she declares that she, not the law, is in charge of how the offender standing before her is to be treated. She is overstepping her office; she is not authorized to take charge over law, but obliged in virtue of her position to remain faithful to law.

This argument does not entail that officials should never exercise positive mercy; in special circumstances, it may be the case that an official should, all-things-considered, overstep her office. The argument does, however, suggest that officials’ obligation to uphold the rule of law gives them standing reason to withhold rather than exercise positive mercy.

The case against positive mercy only gets worse when we recall how positive mercy differs from equity. An act of equity can be defended as an act that derogates from the letter of a law, but upholds the spirit of a law. The law cannot take all particularities into account, but if it could, then it would make the exceptions that the equitable judge makes. In exercising equity, an official does not fail to uphold the spirit of a law; she remains faithful to the spirit of the law and respects it as a constraint on her power. But the same cannot be said of positive mercy. A person who grants positive mercy does not just derogate from a law, she indicts a law as unjust, as a cancer in the system. She takes a noncompliant even insubordinate stance against the law *per se*; as Bentham would put

it, she does not uphold or interpret law, she *over-rules* law. Such conduct may be fit for a legislator (or an autocrat), but certainly not for an official charged to uphold the rule of law.¹⁵⁰

The Strategy of My Response

My response proceeds on two levels. First, consider a political community that has incorporated a practice of mercy into its legal system (I am arguing that political communities have good reason to do this). Certain officials would be legally authorized to derogate from criminal sentencing laws when such laws are unjust and prescribe too much punishment. When an official shows positive mercy in such a system, she may well derogate from *a* criminal sentencing law, but she does not derogate from *the* law. She is not engaging in an act of insubordination against the rule of law; the law allows her action. The merciful official is no more “insubordinate” than for example, a conscientious objector in a political community that makes legal room for conscientious objection. If a legal system allows positive mercy and an act of positive mercy falls within that allowance, then it does not violate officials’ duty to uphold the rule of law.

¹⁵⁰ As Bentham puts it:

To attribute to the legislator a will which you suppose him to have failed of entertaining through...wrong judgment or perverse affections, and to act accordingly, is not to interpret the law, but to act against it: which, in a judge or other officer invested with powers of a public nature, is as much as to over-rule it. The executive power thus exercised approaches to the nature of autocratic power in as far as the effects of it, when exercised, are confined to the individual case in which it is exercised: to that of legislative power in as far as the exercise of it in that instance comes to serve as a rule of decision in subsequent cases which are deemed of the same sort. (*Limits of the Penal Branch of Jurisprudence*, Sec. 16, para. 130, p. 75, in *Oxford Collected Works of Jeremy Bentham*, ed. Schoefield)

The response at this level, however, misses the heart of the problem. The response shows that *in legal systems that have incorporated a practice of mercy*, acts of mercy do not violate officials' obligation to uphold the rule of law. But the response does not show that legal systems *can* incorporate a practice of mercy without posing a significant (perhaps paralyzing) threat to the rule of law in the system. Officials in legal systems are obligated to maintain law's rule and so, at least according to Fuller's eighth principle, they are obligated to keep their behavior congruent with laws as they are, not as they might be. A practice of positive mercy clashes with this obligation because it authorizes officials to derogate from actual laws on grounds that such laws are not what they ought to be. A practice of mercy just does not fit in the context of a system that obliges its officials to remain faithful to upholding laws as they are. To successfully respond to the case against positive mercy at its deepest level, I must show that this clash is merely apparent.

The strategy of my response is to show that contrary to appearance, positive mercy is in the service of the rule of law. I do not deny that officials are obligated to uphold the rule of law. Rather, I affirm their obligation to uphold the rule of law but I argue that this obligation furnishes a standing reason to exercise rather than withhold positive mercy.

My response hinges on an analysis of the rule of law in the criminal justice system. I follow Waldron in his keen recognition that a) the rule of law has to do with the law being "in charge," and his honest admission that b) there is deep debate over how to structure our institutions such that law actually is in charge. Fuller's list of eight principles is not an authoritative and detailed "how-to" guide on how to structure

government such that law is in charge.¹⁵¹ Even if his list were the best guide for structuring the government's basic framework, it may not be the best guide for structuring each particular institution of the government. Indeed Fuller insists that his eight principles are not a good guide for structuring military institutions. How, then, do we figure out the best way to structure a government institution such that law is, if at all, in charge?

My methodological suggestion is that we think about what is valuable about law's being in charge. Then we should aim to structure a government institution in a way that best realizes these values (if they are indeed worth pursuing in the given institution). We should be open to the possibility that the best way to realize these values may not be the same in each institution. There may be no general recipe for having law be in charge of an institution. This thought may lie behind Waldron's suggestion that "the rule of law" is a contested concept; perhaps what it takes for law to be in charge is very different in one institution compared to the next (what it takes for law to be in charge of the criminal justice system may be different from what it takes for law to be in charge of the civil court or the tax system).

To apply this method, I return to the four rule of law values I described above. I argue that the best way of structuring the criminal justice to realize these values is not by obliging officials to sentence offenders in strict congruence with the law. Although adherence to Fuller's eighth principle may be a good, even essential, way to realize rule of law values in some government institutions, I will show that it is not a good way to realize rule of law values in the criminal justice system. I will demonstrate that rule of

¹⁵¹ Waldron, "Rule of Law an Essentially Contested Concept."

law values are better realized in the criminal justice system if judges derogate from, rather than remain congruent with, unjust sentencing laws that prescribe excessive punishment. Or in other words, rule of law values are better realized in the criminal justice system insofar as judges exercise positive mercy.¹⁵²

Given limitations of scope, I cannot analyze the full structure of the criminal justice system. Instead of considering whether rule of law values are best served in a system where police, prosecutors, parole boards, juries and judges exercise positive mercy, I focus only on whether rule of laws values are best served by judges exercising positive mercy.¹⁵³ This is a significant limitation, but it is not “an easy way out.” Compared to other officials, judges arguably have the most robust obligation to uphold the rule of law and retain fidelity to the law. They are, if you will, closest to the reins of law’s rule. The public affords them their position of power on the very strict condition that they do not seize those reins. Judges who are suspected of even mildly deviating from the law in their decisions are harshly criticized as “legislating from the bench,” and

¹⁵² Finnis shares the thought that rule of law values may be best served by officials derogating from laws. He does not, however, join me in stepping into the minefield of explaining how and when such occasions arise.

Sometimes the values to be secured by the genuine rule of law and authentic constitutional government are best served by departing, temporarily but perhaps drastically, from law and constitution. Since such occasions call for that awesome responsibility and most measured practical reasonableness which we call statesmanship, one should say nothing that might appear to be a key to identifying the occasion or as guide to acting on it. (Finnis, *Natural Law and Natural Rights*, 275).

¹⁵³ I suspect that my analysis of a judge’s obligation to follow and apply law can be adapted to apply to these roles. Kadish and Kadish’s consideration of the question in *Discretion to Disobey* seems to support my suspicion. I leave this project, however, for future work.

as Bentham put it, “contriving to steal power.”¹⁵⁴ I argue that such criticism should not be leveled against judges who exercise positive mercy. Judges who exercise positive mercy are not stealing power, but rather fulfilling their obligation to uphold law’s rule as it is best understood in the context of the criminal justice system.

Positive Mercy in Service of Law’s Rule

In what follows, I analyze the four values of the rule of law. I argue that derogation from law in the form of positive mercy rather than strict congruence with law is the best way for judges to instantiate these rule of law values in the criminal justice system.

Reciprocity: The rule of law has non-instrumental value in structuring and maintaining a relation of reciprocity between the ruling and the ruled. The government expects members of the political community to follow the laws and in turn each member has the assurance that the government will follow those laws with respect to them. This is the general idea, but the devil (or rather, a god of mercy) is in the details. The government cannot legitimately demand members of the political community to follow every law. As the practice and study of civil disobedience over the past decades has shown, members of a political community are not required to obey unjust laws.¹⁵⁵

¹⁵⁴ Bentham was particularly critical of derogation in the form of pardon: “from pardon power unrestricted, comes impunity to delinquency in all shapes: from impunity to delinquency in all shapes, impunity to malfeasance in all shapes: from impunity to malfeasance in all shapes, dissolution of government: from dissolution of government, dissolution of political society,” (*Principles of Penal Law*, Part II, Bk. VI, Appendix, p. 530).

¹⁵⁵ See Martin Luther King, “Letter from a Birmingham Jail.” Among others, Rawls, Dworkin, Finnis, and Raz agree that individuals are not morally required to obey unjust laws. Fortunately, it is hard to find exceptions to this trend in the literature.

Respect for the rule of law calls for people to be very careful and conscientious in their consideration of whether a law is unjust. If a law is unjust, however, people are not required to obey it.

With the idea of civil disobedience in mind, let us return to reciprocity and the rule of law. The judge who exercises positive mercy in derogation from an unjust law mirrors the individual who engages in civil disobedience of an unjust law. Each faces a situation in which the constraint of the law misses the moral mark. Reciprocity calls for the constraint to be qualified on *both* sides, not just on the non-official's side of the power relationship. The qualification on the official side might be significantly different and more restrictive than it is on the non-official side. I do not give a general analysis of the qualification; I limit my remarks to the exercise of positive mercy.

Suppose a judge is in a situation such that following an unjust sentencing law would require her to give an offender an excessively harsh sentence. If she withholds positive mercy and gives the offender the excessively harsh sentence, she says, in effect, "The government should not have passed this unjust law. But you are the offender here, so we expect you to bear the burden of your mistake *and our mistake* by serving the excessive sentence that our unjust law requires." Reciprocity in the power relationship is jeopardized, not instantiated by, such a decision to withhold positive mercy. We generally characterize a relationship as abusive, not reciprocal, when the more powerful party makes mistakes and forces the less powerful party to bear the burden of them. The judge who withholds positive mercy sounds like a bit like an abusive father who tells the child he beats, "Just look what you have made me made me do!" That kind of behavior is not appropriate in a reciprocal relationship.

A party in a reciprocal relation recognizes her own faults and shoulders the burden of them, she does not try to pass them off as the other's faults and force the other to suffer for them. The best way to maintain this kind of reciprocal relation in the criminal justice system is not for judges to act in accord with every constraint the law places on them. Rather, the best way to maintain reciprocity is for judges to violate the constraints of unjust sentencing laws by exercising positive mercy.

Autonomy: Although the value of reciprocity is best served through the judicial exercise of positive mercy, the other three rule of law values may be best served by judges withholding positive mercy. Is respect for individual autonomy best served by judges withholding or exercising positive mercy?

In order for individuals to exercise their capacity to make meaningful decisions that guide their own lives, they need to be able to form reliable expectations about what behavior the political community will and will not punish. The laws—as long as they are clear, stable, non-retroactive, and consistently enforced—allow people to form stable expectations about the kind of legal treatment they will receive. When officials judge an individual on the basis of such laws, the individual is judged on the basis of standards that she knew (or could have known) about and could have followed. She is judged as a self-directed, responsible agent. Any punishment she receives is a direct result of her own choices, not a result of the passing whim of an official.

When a judge punishes an individual in derogation from law, the individual's autonomy might be threatened. If a judge punishes an individual *more* than the law allows, then the individual's autonomy is likely threatened. She is judged on the basis of a standard (the judge's) that she did not have a reasonable opportunity to follow. She

knew or could have known about the law, but she did not know and could not have known about the judge's standard. If she had known about the judge's standard, then she might have chosen not to commit the crime so as to avoid this added punishment. But she was not given this opportunity. The added punishment is not attributable to her own choices. The judge who gives her such punishment is not treating her as a self-directed autonomous agent.¹⁵⁶

When a judge punishes an individual *less* than the law prescribes, however, the judge does not necessarily fail to respect her autonomy.¹⁵⁷ To the extent that the individual receives any punishment from the judge, the punishment is still attributable to her choice to commit the crime. If she had known about the judge's more lenient standard, this knowledge would almost surely not have prevented her from choosing to commit the crime. She violated a law that she could have obeyed and that she knew could have come with *at least* as much punishment as the judge makes her suffer. One cannot legitimately object that the judge has failed to respect her as a self-directing agent.

An offender like Oliver Wendell Holmes' "bad man" might try to object as follows. Had she known the more lenient standard she would be judged by, she would have chosen to commit more (or worse) crimes. The discrepancy between the judge's decision and the law hinders her ability to form reasonable expectations about exactly

¹⁵⁶ This line of reasoning is often used to explain why retroactive laws are inconsistent with the rule of law.

¹⁵⁷ One might worry that there are exceptions here. On Herbert Morris's view, if a judge punishes an offender less based on a false judgment that the offender is incapable of self-directed action (when the offender actually is capable), then the judge fails to respect the offender's autonomy ("Persons and Punishment"). But such a decision would not qualify as positive mercy because it does not give the offender substantively just treatment.

how much mischief she can get away with without suffering punishment (or some minimal amount of it). The judge is hindering her ability to make efficient crime choices that are important for guiding her life and so the judge is hindering her autonomy.

This is not a legitimate objection. The objector seems to assume that respect for an individual's autonomy requires supplying her with information to make *all* of her options clear. But that assumption is false. Respect for autonomy (in the valuable sense that the rule of law is interested in protecting) means respect for a person as a responsible, self-directing agent. To respect a person as a self-directing agent, the law must give her a reasonable opportunity to avoid punishment, but it need not give her exact assurance as to what will result from every legal risk she takes. To insist that the law must provide such assurance out of respect for autonomy is to confuse autonomy with the maximization of bare negative liberty. I maintain that although respect for autonomy is a value of the rule of law, maximizing bare negative liberty is not.¹⁵⁸

I have argued that respect for individual autonomy does not necessarily count against positive mercy. Now I turn to argue that positive mercy actually helps the rule of law in instantiating respect for autonomy. I assume, following the general lead of Herbert Morris, that political communities can morally justify some amount of punishment to some individuals by arguing that punishment is necessary in order to respect them (and others in the community) as self-directed agents.¹⁵⁹ When a judge withholds positive mercy when she has occasion to exercise it, she gives an offender an

¹⁵⁸ I stand with Rawls on this point and not with Hayek. See *supra* note 143 for more discussion.

¹⁵⁹ Morris, "Persons and Punishment."

excessive punishment that goes above and beyond the amount needed to respect the offender as a self-directed agent. This additional punishment is not innocently gratuitous from the point of view of autonomy. On the contrary, it is generally damaging to the offender's autonomy. As I explain more fully in Chapter 5, most modern punishments take the form of incarceration that not only severely diminishes an individual's opportunity to make choices that guide her life, but can also degrade an individual's very capacity to make choices for herself (by both lack of exercise and by jeopardizing the individual's self-respect). By giving just rather than excessive punishment, the exercise of positive mercy avoids this outcome. That is, positive mercy respects rather than jeopardizes the autonomy of the offenders who receive it.

A tool to guide behavior: The rule of law has instrumental value as a tool for social coordination. The government uses laws to achieve various goals (such as environmental protection, education, maintenance of security, etc.). For law to achieve such goals, it must be capable of guiding human behavior. According to Raz, law guides behavior by playing a role in people's practical reason.¹⁶⁰ In order to play such a role, the law must (among other things) be clear, stable, prospective, and general. Also, those enforcing the law must not deprive the law of its ability to guide human behavior. Officials must not "pervert" or "subvert" the features of law (clarity, stability, etc.) that are needed for the law to play a role in people's practical reason. When officials derogate from law, they tend to subvert law's clarity and thereby tend to diminish law's ability to play a role in people's practical reason. This means that law will be less effective in

¹⁶⁰ I draw heavily here on Raz's discussion of the rule of law in *The Authority of Law*, 210-229.

guiding people's behavior and in turn be less effective at achieving its variety of purposes. The deterioration of law's ability to achieve its purposes is deleterious for the political community insofar as those purposes are just.

There are two ways that positive mercy could exert an effect on law's ability to guide human behavior: it could have an effect on the law's general ability to guide human behavior or on particular laws' ability to guide human behavior. I consider these effects in turn.

Law's general ability to guide human behavior depends on the public's trust that the law will be enforced reliably. In strong, well-established legal systems, people's trust in law enforcement will almost surely not be shaken by the judicial exercise of positive mercy. A public with reasonable trust in its system will not infer that positive mercy will lead to a scourge of judges and other officials off-roading from law. Of course it is possible for them to draw such an inference, but it is highly unlikely in practice. The scope of mercy is very limited; it is not general judicial discretion to disobey any laws. Judicial mercy is limited to derogation from sentencing laws. A judge can only exercise positive mercy in the determination of sentences, not in the determination of guilt/innocence for crime. Offenders who are granted judicial mercy are still arrested, put on trial, and punished justly. Given mercy's limited scope, it is highly unlikely that the exercise of positive mercy will shake public trust in law enforcement. As such, we should not expect exercises of positive mercy to seriously threaten law's general ability to guide human behavior. The possibility may be more of a live concern in a fledgling legal system that has not yet won public trust in its ability to enforce clear, stable, and promulgated rules. But in most modern political communities, this is not a realistic

worry;¹⁶¹ as Kennedy puts it, “a people confident in its laws and institutions should not be ashamed of mercy.”¹⁶²

In some contexts, the exercise of positive mercy may actually enhance law’s general ability to guide behavior. If, as is arguably the case in America, there are unjust sentencing laws that disproportionately affect certain neighborhoods or communities (like inner city ghettos), then people in those neighborhoods may increasingly view the legal institution with justified derision and suspicion. The police have a more difficult time enforcing law in these neighborhoods (those who help the police are stigmatized as “snitches”) and so law loses some of its ability to guide behavior. In even worse circumstances, large proportions of neighborhoods can be incarcerated due in part to unjust sentencing laws. When this happens, there is often an inversion of value within the neighborhood;¹⁶³ illegal behavior and incarceration ceases to be seen as something to avoid and is instead seen as a badge of honor.¹⁶⁴ In such circumstances, law has certainly lost its ability to guide human behavior in the right direction. The exercise of positive mercy can help the law gain back its ability to guide behavior in these neighborhoods in two ways. First, by decreasing the proportion of the population that is incarcerated. Second, by showing that the criminal justice system is trying to do substantive justice rather than continuing to enforce unjust rules. In these ways, the exercise of positive

¹⁶¹ I maintain that the worry is far-fetched. If my opposition thinks it is a live concern, then the burden of proof is on her to demonstrate it.

¹⁶² Address at the American Bar Association Annual Meeting.

¹⁶³ Shelby makes a similar point in “Justice, Self-Respect, and the Culture of Poverty.”

¹⁶⁴ See, for example, Oliver, “‘The Streets’: An Alternative Black Male Socialization Institution.”

mercy might help win back at least a measure of public trust in the legal institution in these neighborhoods. Even if positive mercy does not win back any public trust, it will at least not further alienate trust. In this respect, it does better than the status quo of withholding positive mercy.

Positive mercy does not hinder and may even help law's general ability to guide human behavior. But does it hinder particular laws' ability to guide behavior? The exercise of positive mercy might affect law's ability to enter into an individual's reasoning over whether to obey particular criminal laws.¹⁶⁵ With mercy in the picture, people (like Holmes' bad man described above) have to take one more factor into account in their decisions about whether to commit crime. The factor is: will I get a reduced sentence through positive mercy? One might think that this added "mercy factor" diminishes law's ability to guide behavior because it makes the consequences of violating law less clear. But does it? The "mercy factor" certainly does not inject any additional uncertainty about the consequences of violating laws in all cases. If the sentencing law covering a given offense is not unjust, then there is no chance of positive mercy (unless the judge makes an error or abuses her power—but this is a possibility with or without positive mercy). If the sentencing law is unjust, then the "mercy factor" does introduce a

¹⁶⁵ The worry here is potentially overstated. A general practice of positive mercy may have little effect on diminishing predictability. If judges abuse their power to show positive mercy, then predictability might diminish. But if judges use their power to show positive mercy publically and in all and only the appropriate cases (i.e. by giving just treatment when unjust laws prescribe excessively harsh treatment), then predictability might only diminish marginally. Insofar as members of the political community can spot an unjust law as well as a judge, they will be able to predict when positive mercy will be shown. They might not be able to pinpoint what the just treatment will be in positive mercy cases, but they should have a general idea of the range of options. They will be no less able to predict the sentence in a mercy case than in a usual case where a judge is given sentencing discretion.

degree of additional uncertainty about the consequences of violating the law. The judge may or may not be merciful (as I discuss later, there is no way to force judges to be merciful) and the agent has to take this uncertainty into account. Positive mercy marginally decreases the law's ability to influence practical reason in these cases, but the decrease here is marginal and well-worth accepting.

Suppose there were no positive mercy. Then an agent considering a given crime would know full well that if the sentencing law covering her offense were unjust, she would receive an excessively harsh sentence. There would be no added uncertainty about the sentence due to a "mercy factor." So law would be marginally more effective at guiding practical reason. But there is no intrinsic value in this marginal improvement. One might argue that the marginal improvement has instrumental value. The agent considering crime may be less likely to commit the crime because she has greater confidence that she will be harshly punished. Without positive mercy, the law might be more effective at deterring crime.

But this argument fails to fully appreciate that the sentencing law is *unjust*. Even if it does deter crime, it does so in an unjust fashion. An example helps explain this point. Consider the three-strikes law in California. Suppose a person has committed two offenses that count as "strikes." She knows that if she is found guilty of a single petty offense, she will get a twenty-five to life sentence. This knowledge might deter her from committing a third crime, but it will also make her live the rest of her life in fear of committing some third offense (being in the wrong place at the wrong time, having a minor drug relapse, or being set-up in some way). With this disproportional threat constantly hanging over her, she cannot put the past fully behind her or say she has fully

paid her debt to society. The sentencing law may deter, but because the way in which it deters is unjust, the deterrence is not to be celebrated as a good (or even an innocent) achievement. The fact that unjust sentencing laws are marginally better at influencing people's practical reason in an unjust fashion is not a triumph for the rule of law. It is almost always an embarrassment.¹⁶⁶ Positive mercy helps the rule of law *avoid* this embarrassment.

Transparency: The rule of law has instrumental value in keeping the use of power transparent to the public. When law rules, people know (or can know) that power will be kept within the constraints of law, for better or for worse.¹⁶⁷ This transparency affords the law the predictability and clarity needed for it to play a role in people's practical reason, but it also has two other instrumental values. First, it helps prevent the elite from gaining an unfair advantage. Second, it gives people the knowledge that is required for them to engage in just reform of the law. I consider these values in turn. I show that mercy does not obfuscate transparency in a way that unfairly favors the elite. Also, mercy helps rather than hinders the people's ability to engage in just reform.

When the workings of power are secret, the elite tend to get an upper hand because they can often gain access to the secrets and use them to their advantage. For

¹⁶⁶ There may be a handful of unjust sentencing laws that a) have a defect of justice that makes them determine substantively unjust outcomes in recognizable types of cases, but b) deter in a just fashion and c) would lose justifiable deterrent value if their defect were removed. Positive mercy with respect to these laws might marginally diminish the ability of these particular laws to guide behavior. The loss of value in such cases, however, is generally outweighed by the law's increased ability to achieve another purpose: the determination of substantively just outcomes.

¹⁶⁷ Lon Fuller, among others, endorses this value of rule of law. See Postema, "Law's Ethos."

example, those close to the King know the “magic words” to find favor with the King (or, to use a more modern example, those close to legal officials know just how to bribe them in order to get legal favors). But there is no secret way to find favor with a judge if she is bound to always apply the law and never derogate from it. Allowing judges room to derogate from laws *in general* seems to allow room for “secret rules” on how to win the judge’s favor. Allowing judges to derogate from laws to show positive mercy in particular, however, poses no such threat. The judge can only grant positive mercy if the relevant sentencing law is unjust. The elite may be able to hire better lawyers to argue that certain sentencing laws are unjust, but their doing so is not pernicious. They are not “kissing the king’s ring” to win favors for themselves and not others. Rather, they are providing an important public service. If the elite’s lawyers decisively reveal that sentencing laws are unjust and this knowledge is public, then society as a whole is better off for knowing that a cancerous law is present in their system. That knowledge, although it may initially be applied primarily to the elite, will also benefit the less fortunate if and when they are subject to the same laws. Unlike secrets about how to win favor, public knowledge that a law is unjust is, at least in principle, transferable from the elite to the less fortunate. Positive mercy does not threaten transparency in a way that necessarily gives the elite an upper hand.

Transparency is not only valuable as an equalizer, but also as a necessary tool for just reform. If people know what the laws are, they can take an active and critical stance on their laws. They can engage in protest against unjust laws (or any laws for that matter) and make it difficult for the powerful to maintain unjust laws. Even if people cannot exert enough influence to make changes in unjust laws, transparency at least

affords them the opportunity to engage in symbolic protest or opposition. If judges derogate from the promulgated laws in a willy-nilly but significant fashion, then people will not know what the real rules in their community are. They cannot take a stance against rules that they do not know exist. If judges, for example, lie whenever they confront an unjust law in the system, then people will not be aware that the law needs to be changed.¹⁶⁸

Although the reform-value of maintaining transparency is jeopardized by general and widespread judicial derogation from law, it does not count against the particular kind of derogation that is involved in exercising positive mercy. Positive mercy does not necessarily obfuscate the fact that there are unjust laws in the criminal justice system. A judge can exercise positive mercy in a public fashion. When a judge publically exercises positive mercy, she communicates the fact that the relevant sentencing law in the case before her is unjust. Her act is open to the eye of the public; it signals rather than hides to the public the need for reform. Of course there is no guarantee that judges will exercise mercy in a public fashion. They might show mercy in a secretive, deceptive, or behind-the-scenes fashion. This potential problem, however, is not a reason for judges to withhold positive mercy. Rather, it is a reason for judges to show positive mercy in a public rather than a secretive fashion. Having a general practice of mercy in a criminal justice system can help encourage judges to make their exercises of mercy public.

¹⁶⁸ Dworkin suggests that judges may face the difficult normative decision to either lie or resign in cases where the law is unjust (*Taking Rights Seriously*, 327, 342). With respect to unjust sentencing laws, there is a third and morally superior option here: exercise positive mercy.

In addition, a general practice of positive mercy can play a role in reforming unjust laws from *within* the legal institution. If a judge exercises positive mercy in a given case, then other judges facing similar cases can follow her lead. Sometimes these exercises of positive mercy can add up to a legal precedent that might effectively change the law. There is a great deal of value to building-in mechanisms for reform of the law from within the legal institution.¹⁶⁹ Law's rule, like almost any kind of rule that aims at stability and justice over time, is generally better if it has mechanisms for self-correction. A general practice of positive mercy within the criminal justice system provides the rule of law with such a mechanism.¹⁷⁰

Both the social equalizing effect and the opportunity for reform that are afforded by transparency in law's rule are best served by the exercise of positive mercy. I have also shown that these values are best served by the *public* exercise of positive mercy that is preferably embedded in a general practice of mercy.

I have argued that the exercise of positive mercy helps rather than hinders the instantiation of the four rule of law values. The best way for law to rule in the context of criminal justice systems is not by insisting that judges strictly adhere to Fuller's eight principles. Although rule of law values are *generally* best served by congruence between

¹⁶⁹ Hampton, "Democracy and the Rule of Law," and Valcke, "Civil Disobedience and the Rule of Law."

¹⁷⁰ A passage in Seneca echoes this point. He writes that a person trained to grant mercy says to herself, "I keep mercy always in readiness; such guard do I keep over myself as though I am going to render an account of the laws that I have called forth from the darkness of decay into daylight," ("On Mercy," in *Dialogues and Essays*, trans. Davie, 188-189). Mercy performs a critique of the laws (and other social rules), an "accounting" of them, which shines the light of justice upon them and openly reveals any decay in them. A merciful judgment remedies the decay not only by giving a just verdict in an individual case, but also by "calling for[th]" improvement in the laws themselves."

the law and judicial decisions, this is not so in the particular context of the criminal justice system. In this context, rule of law values are best served by selective incongruence between the law and judicial decisions. The values are best served when judges exercise positive mercy in derogation from law rather than withholding positive mercy in keeping with the law. The best way to structure a criminal justice system for law's rule is therefore to have judges practice positive mercy. In this sense, positive mercy is a service to law's rule, not a mutiny against it.

A judge who withholds positive mercy on the purported grounds of serving the rule of law is committing an error in most cases. The judge is not demonstrating healthy fidelity to the rule of law, but rather a faith in the fact that Fuller's eighth principle is the way to maintain the rule of law. I have argued that such faith is misguided in this context. A judge who withholds positive mercy on such grounds may not be vicious, but in most cases she is mistaken about what her obligation to uphold the rule of law means and entails in the particular context of criminal justice. Her mistake here can be labeled as improper *formalism*. She may be accustomed to hearing, saying, and thinking that her obligation to uphold the rule of law means that she has "no choice" but to follow the law. This mantra may be useful in most contexts, but it may make her blind to the fact that her obligation to uphold the rule of law *does* give her a choice on whether to follow unjust sentencing laws. In fact, her obligation to uphold the rule of law generally stands as a reason to choose in favor of derogating from the law to show positive mercy.

Sometimes a judge who withholds positive mercy on purported rule of law grounds is not just engaging in misguided formalism, but a much worse failing. Throughout history, government officials have often abused the idea of "rule of law" as a

guise to “justify” cruelty and brutal injustice. Consider a judge who knows that the rule of law, properly understood, stands in favor of mercy. Suppose she nevertheless withholds positive mercy when she should exercise it. She proclaims that she “must” apply the law and that the law gives her “no choice” but to sentence the offender to an excessive punishment. She not only abuses the offender, but also abuses the idea of the rule of law.¹⁷¹ By doing willful violence to the offender and the rule of law, she is guilty of the cruel formalism that has often greased the wheel for injustice under the guise of legality.¹⁷²

The judge who stands willing to show positive mercy does not engage in either a misguided or a cruel formalism. She has a healthy understanding of the contours of her role in governing people by law. To use the words of David Lyons, she is not playing some esoteric game under “the naive assumption that ‘the proper function of a court is to apply an established rule of law to the dispute before it.’ [S]he would do more justice to the judicial role.”¹⁷³

¹⁷¹ Judith Shklar cynically argues that abuse like this has been so extensive that it has rendered the rule of law nothing more than “a bit of ruling class chatter... ‘the Rule of Law’ has become meaningless thanks to ideological abuse and general overuse,” (“Political Theory and the Rule of Law,” 21).

¹⁷² For insightful discussion of how the rule of law has been used and abused see Dyzenhaus, *Hard Cases in Wicked Legal Systems*.

¹⁷³ “Justification and Judicial Responsibility,” 198. Lyons’ work in this paper paves the way for my own. He argues that “if a decision cannot be justified on its merits, then there must be a morally adequate argument for following the rule, or else the decision cannot be justified at all,” (ibid.). I have argued that with when occasions for positive mercy arise, the decision to give an offender excessive punishment cannot be justified on its merits nor is there a morally adequate argument for following the rule.

I have argued that judges have reason to exercise positive mercy—that is giving offenders just punishment when unjust laws (or other unjust social rules) call for (or even require) them to give harsher punishment.¹⁷⁴ The argument, at least with respect to outlaw positive mercy, is not that judges have reason to do substantive justice in individual cases and that this reason sometimes outweighs the reason to do uphold the rule of law (or, in roughly similar terms, to do procedural or formal justice).¹⁷⁵ I have argued that positive mercy stands in service to the rule of law, not against it. A judge’s decision to exercise positive mercy is generally not a decision *between* upholding the rule of law and doing substantive justice in a given case. The judge does not have to compromise one of these for the other. *Both* serving the rule of law and delivering substantive justice stand in favor of the judicial exercise of positive mercy. There may, as always, be special considerations that complicate individual cases, but as a matter of general guidance, judges have standing reason to show positive mercy in order to do justice under the rule of law.

Perhaps, however, I have over-simplified the matter by using “exercise positive mercy” as a success term. If a judge derogates from a just law and gives an offender an unjust punishment (be it too much or too little), she has not exercised positive mercy in

¹⁷⁴ My general definition of positive mercy has to do with harsh treatment in general, not necessarily punishment. In the context of criminal justice, however, the relevant kind of harsh treatment is generally punishment.

¹⁷⁵ Rawls suggests such an argument: “Sometimes we may be forced to allow certain breaches of the precepts of legality if we are to mitigate the loss of freedom from social evils that cannot be removed, and to aim for the least injustice that conditions allow,” (*Theory of Justice*, 243). I have no objection to an argument of this structure in other contexts, but I think we can do better than this in the context of criminal justice with respect to unjustly harsh sentencing laws.

my sense. If a political community incorporates a general practice of mercy that authorizes judges to exercise positive mercy, then this does not mean that judges will use that authorization appropriately. They may use their power not to exercise positive mercy, but rather to decrease sentences in an unjust manner. Is this a risk worth taking? Should political communities incorporate a general practice of mercy just because judges have reason to exercise positive mercy in certain cases?

I do not answer this question; it is a question that is best answered by doing a more substantive policy analysis in the context of a particular political community. I instead argue that political communities have *a* reason to adopt a general practice of positive mercy in the law-abiding fashion for all officials in the criminal justice system and a general practice of positive mercy in the outlaw fashion for (at least) judges in the criminal justice system. The reason to adopt a general practice is that doing so helps encourage the judicial exercise of positive mercy that I have defended in this section. A general practice particularly encourages judges to exercise positive mercy in a public fashion that I have shown to be especially valuable. Upon further policy analysis, it may turn out that the reason in favor of adopting a practice of judicial positive mercy is outweighed by other considerations in a particular criminal justice system. I hope to have shown, however, that we at least have reason to conduct this analysis. Positive mercy deserves a place on the menu of options about “what is to be done” in response to injustice in criminal justice systems.

Section IV: An Objection

A deeper objection suggests that positive mercy does not even deserve to be considered as an option in modern political communities (at least not liberal ones). The

objection presses on the fact that the concept of positive mercy is parasitic on the idea that certain treatment of criminal offenders is substantively just and other treatment is not. The idea of judges making extra-legal determinations of substantive justice often makes audiences in liberal, pluralistic communities legitimately concerned. Substantive justice with respect to criminal justice from one comprehensive moral view can look quite different from substantive justice from another comprehensive moral view. Instead of trying to decisively determine which substantive view is “right,” liberal political communities aim for procedural justice. They have just procedures that allow the community to settle on certain answers about what qualifies as justice for the purpose of living together as a political community. When a judge exercises positive mercy, she appears to step outside of this framework. She self-righteously deems that a law which the democratic legislature has passed is unjust—not unjust in the way it was passed, but because it is substantively unjust. As Waldron puts it, even if a judge’s dissent from law “is conscientious and based on impeccable moral arguments...it is announcing in effect that it is better to revert to a situation in which each acts on their own judgment about justice.”¹⁷⁶ Liberal political communities have by and large decided that we should avoid reverting to such a situation.

The way to address this concern is not to take positive mercy off the table in liberal political communities. I do not dispute the fact that members of these political communities disagree about a wide variety of questions surrounding substantive criminal

¹⁷⁶ *The Dignity of Legislation*, 59. As Postema puts it, “Kant taught well the lesson that exercising power with the most sincere conviction...of right in the absence of public accountability is, from the perspective of one subjected to that exercise of power indistinguishable from being subjected to action on whim,” (“Law’s Ethos,” 1864).

justice. I also agree that the passage of law through procedurally just channels is a very important way for members of a political community to settle those disagreements for the purpose of living together. We should not forget, however, that despite wide disagreement on matters of substantive justice, a political community might still agree on a few matters of substantive *in*justice. Even if they disagree about what the justifiable ends of criminal punishment are and how they are to be accomplished, they might still have an “overlapping consensus” on which punishments are unjust. Consider a punishment in a given case that goes above and beyond what is required to achieve any and all of the justifiable ends of punishment that public reason recognizes. Insofar as people modulate their judgments about punishment in accord with public reason, they will agree that the punishment is unjust.

It would be naïve to hold, however, that the legislature could never pass laws that require such punishment (or that it would always remove existing laws that require such punishment). If there is a law that requires such an unjust punishment and the judge exercises positive mercy, she is not turning her back on a shared view of justice in favor of her own personal conception of justice. She is rather defending the shared view of injustice.¹⁷⁷

¹⁷⁷ One might think that American jurisprudence surrounding the Eighth Amendment on cruel and unusual punishment works in such a manner. I think this is an open question. The court often determines whether punishment is “cruel and unusual” on the basis of legal precedent and not on the basis of considerations of substantive justice (for example, in *Lockyer v. Andrade* the court’s decision was based on the thought that it had no choice but to follow the law, see Pater, “Struck out Looking”). Sometimes they consider the public’s “evolving standard of decency,” but again it is not clear that this tracks public reason about criminal punishment or public fears about crime and the political rhetoric that leverages these fears. I leave further investigation of this question for future work.

Although the objection does not necessarily preclude positive mercy, it does raise an important concern about it. The right way to address this concern is to develop a public accountability scheme for the judicial exercise of positive mercy. As discussed in Section I, the political community can set up a way to review the judicial exercise of mercy powers. As I mentioned previously, such a scheme can sanction judges who try to pass off nepotism or racism as “positive mercy.” But it can also play another role. A review system can examine whether judicial use of mercy power is consistent with the shared conception of substantive justice in the political community. The review system would encourage members of the political community to engage in meaningful dialogue about what they agree is substantively unjust treatment. It could take the results of such dialogue and apply them as follows.

Suppose the review finds that a judge showed mercy in a case where the law required a punishment could not be justified by any proper use of shared public reason. In this case, the judge’s decision to show mercy should easily pass review. Other cases will be more difficult. A law may require punishment that appears unjustly harsh but which might also be consistent with a shared view of justice in the political community. If a judge exercises positive mercy in these circumstances, then she takes a risk that “tests the water.” Her decision may, upon review, be deemed inconsistent with the shared view of substantive justice in the political community. If she made her decision *conscientiously*, however, she should not be sanctioned for her error. Her decision should simply be overturned. In a third kind of case, it may be reasonably clear that a punishment is consistent with the shared view of justice in the political community. If a judge exercises positive mercy in these circumstances, then she takes a grave risk. Her

decision should not only be overturned, but she could be sanctioned legitimately for her behavior. Unless she can demonstrate that the shared view of justice is deeply pernicious, then the political community is probably right to sanction her for trying to substitute her own personal conception for the shared conception of justice in the law.

By having such a review scheme, judges will think twice about their decisions to show positive mercy. Are they inserting their own personal conception of substantive justice into the law or are they acting in service of the public conception of substantive justice? As the practice and study of civil disobedience attest to, the fact that this kind of question is *difficult* is not a reason to stop thinking about it and acting on one's best judgment of it. There is considerable reason, however, to encourage and exercise caution in answering this kind of question. Judges should exercise caution in using their power to show positive mercy and an accountability scheme that involves sanctions for reckless use of this power is one way to encourage such caution.¹⁷⁸ The scheme should not stifle positive mercy, however, by assuming from the outset that all sentencing laws are consistent with the public conception of justice. Instead, it should rely on judges to use their best judgment and then, when needed, there can be substantive public debate on whether their judgment was correct (or at least conscientious).

¹⁷⁸ Note the difference here between law-abiding mercy and outlaw mercy. With respect to law-abiding mercy, I suggest that the primary worth of the accountability scheme will be to help encourage officials in using mercy power that might otherwise remain dormant. With respect to outlaw mercy, I suggest that the primary worth of the accountability scheme will be to caution judges from abusing mercy power that might otherwise be used too liberally. In both contexts, the idea of the accountability scheme is to try to achieve the right balance so that mercy powers are neither abused nor neglected, but used to exercise positive mercy.

There is no doubt that positive mercy in a political community ultimately requires its members to trust their judges (and those reviewing them) to make good judgments about substantive justice. I hope to have shown, however, that this trust in a judge's judgment is not to be confused with giving judges the reins over law. I have argued that positive mercy does not jeopardize, but rather best serves law's rule in a political community.

Conclusion of Part I:

Much contemporary discussion of mercy has focused on negative mercy: compassionately sparing an offender from harsh treatment that she deserves. Drawing on Seneca's discussion of mercy, I have articulated a different concept of positive mercy: treating a person justly when social rules call for harsher treatment. Negative mercy has conceptual features that raise doubts about its use in the criminal justice system. Positive mercy lacks these problematic features and, as I have shown in this chapter, it has a great deal of potential in a criminal justice system. Judges have reason to exercise positive mercy insofar as they have reason to seek justice under the rule of law. Given what stands in favor of positive mercy, it deserves a place at the table in discussion of mercy in criminal justice; it would be myopic to give place to negative mercy but not positive mercy.

My purpose in these two chapters is not solely to articulate two concepts of mercy and compare their merit in the criminal justice system. Recall from my introduction that my overarching aim is to investigate what might be done in face of injustice in the criminal justice system. If unjust rules are lurking in a criminal justice system, they should be changed. If they are laws, they should be changed by passing new legislation.

If they are informal rules or institutional policies, they should be reformed using various means that might include legislation. But these changes and reforms do not happen overnight and sometimes we have reason to think they will not be changed anytime in the near future. In the meantime, officials have reason to show positive mercy rather than abide by and uphold these unjust rules.

As a criminal justice system improves its rules, the occasions for officials to show positive mercy diminish. But our rules about punishment are far from perfectly just. As long as our criminal justice systems harbor some unjust rules, we still have good reason to incorporate, or at least further investigate incorporating, positive mercy into these systems. If I have succeeded in my defense of positive mercy, the reader may be curious. Are there unjust rules in her political community's criminal justice system that officials have reason to respond to with positive mercy?

In the next part of the dissertation, I turn to the task of identifying unjust rules that lurk within many modern criminal justice systems. I do not comb through the rules of existing criminal justice systems and pick out the unjust ones. Instead, I raise a red flag about a set of rules that are found in many modern criminal justice systems. I am concerned about rules that regulate the political community's expression of a certain kind of blame – what I call abrasive blame – toward criminal offenders. After identifying abrasive blame and explaining why it is permissible for the political community to express it toward some criminal offenders, I argue that it is not permissible for the political community to express it toward at least two types of offenders: battered and fragile offenders. I raise a red flag about rules that call for officials in the criminal justice system to express abrasive blame toward battered and fragile offenders. We have reason

to think such rules are unjust; they call for impermissible treatment toward a significant set of offenders. If these rules are indeed unjust, they should be changed. In the meantime, officials have reason to treat these offenders justly (however that may be) and withhold the harsh abrasive blame that the unjust rules call for. That is, if the rules about abrasive blame are indeed unjust, officials have reason to exercise positive mercy toward a significant set of offenders.

PART II.
CHAPTER 3: THE STING OF BLAME

“Down, down to hell, and say I send thee thither.”

~Shakespeare¹⁷⁹

“Like Judas of old, you lie and deceive...For threatening my baby, unborn and unnamed,
you ain’t worth the blood that runs in your veins.”

~Bob Dylan¹⁸⁰

* * * * *

“Anger should be circumvented... When Socrates once received a box on the ear, the story goes, he merely said it was a nuisance that men couldn’t tell when to wear a helmet when going for a walk. What matters is not how an offence is delivered, but how it is endured.”

~Seneca¹⁸¹

* * * * *

Sometimes people express attitudes that are meant to sting a person because she did something (or many things) wrong. I call such expressions *abrasive blame*. It is difficult to explain the “sting” that abrasive blame is meant to inflict. I invite the reader to survey her own moral experience of either stinging or being stung by abrasive blame. I provide the first two quotes above to help jog the reader’s memory. I can say what the “sting” is not – it is not a verbal “slap on the wrist” intended only to improve the

¹⁷⁹ *Henry VI*, Act V, Scene 6, line 3064.

¹⁸⁰ “Masters of War.”

¹⁸¹ “On Anger,” in *Dialogues and Essays* trans. Davie, 27.

wrongdoer's future behavior or advance her moral education. The sting of abrasive blame is much more personal than a teaching tool.¹⁸² The sting of abrasive blame is also not merely an attempt to "let off steam" or assert one's rights. A victim with healthy self-respect can let off steam by taking a brisk walk and can assert her rights by saying, "You did something wrong, but I forgive you."¹⁸³ Abrasive blame may involve letting off steam, asserting rights, and morally educating, but it paradigmatically involves a distinct something else: the sting. In short, abrasive blame is not benign.¹⁸⁴ Abrasive blame is undoubtedly part of the human experience. But is abrasive blame morally permissible? It stands at odds with a general moral injunction against intentionally hurting people.

I investigate abrasive blame and its permissibility in this chapter with the ultimate aim of investigating whether the institutional expression of abrasive blame toward offenders in the criminal justice system is morally permissible. Abrasive blame toward offenders in the criminal justice system is often taken for granted in moral critiques of criminal justice. It has not been recognized as a distinct way in which we treat criminals that stands in need of justification. Abrasive blame may be off the critical radar because people have underappreciated its significance or because they assume it is obviously permissible. Over the course of the next three chapters, I will argue that abrasive blame toward criminal offenders does stand in need of justification, that the justification is

¹⁸² I take this point from Peter Strawson, "Freedom and Resentment."

¹⁸³ I defend this claim elsewhere in "Forgiveness as an Alternative Response to Wrongdoing."

¹⁸⁴ Susan Wolf aptly identifies these negative reactive attitudes and calls them "angry blame," ("Blame, Italian Style"). I chose "abrasive blame" to emphasize that this set of reactive attitudes involves the intentional infliction of a sting, something that hurts.

fragile (at best), and that it is prima facie morally wrong for the political community to abrasively blame some offenders. Although it may appear artificial in some ways to draw a distinction between abrasive blame and closely related behaviors (or aspects of our moral psychology), the proof here is in the pudding. I contend that carving out abrasive blame as a category of its own has dividends in a moral critique of the way we treat one another, especially the way we treat criminal offenders who are often the targets of some of our most powerful and hurtful expressions of abrasive blame.

In this chapter, I identify abrasive blame and I dispel the commonly held notion that a person's commission of a wrong action is itself an adequate justification for people to abrasively blame her. The judgment that a person committed a wrong action is a necessary and sufficient reason to motivate abrasive blame. But neither the commission of a wrong nor the judgment of the commission is a sufficient reason to morally justify abrasive blame. The fact that a person commits a wrong within the context of a relationship or community can, however, provide an adequate justification for people in that context to abrasively blame her under the appropriate conditions. I do not argue that abrasive blame is usually morally wrong. As a matter of sociological fact, I suspect most people are usually in personal relationships or communities in which abrasive blame is permissible. For now, however, I leave the sociology to the experts and put the conclusion of my argument in a qualified form: abrasive blame is morally permissible only within the context of certain personal relationships and communities. This conclusion paves the way for my work in Chapters 4 and 5 in which I argue that abrasive blame toward criminal offenders is only justified under certain conditions (that may or may not turn out to be met).

In Section I, I clarify what I mean by abrasive blame. In Section II, I argue that abrasive blame is morally permissible only within the context of certain relationships or communities. This argument calls for two supporting arguments as well as a careful explanation of what I mean by a personal relationship and community. I give these arguments and explanation in Sections III-V. In Section VI, I raise and respond to an objection and contrast my view of blame in personal relationships to a closely related view developed by T.M. Scanlon.

I engage in considerable analysis of the conditions under which abrasive blame is justified in personal relationships for its own interest in Section V, but also for another reason. People may be hesitant to accept my claim that the commission of a wrong action is not sufficient to justify abrasive blame. This hesitancy may be due to the intuition that it “just seems obviously right” for us to abrasively blame certain people in our lives for the sole reason that those people have done something wrong (for example, when a friend has lied, when a partner cheats, when a colleague insults or takes others’ work for granted, etc.). My discussion of abrasive blame explains that this intuition is often right, but it cautions us from importing it to justify the expression of abrasive blame toward anyone.

Section I: Abrasive Blame

Abrasive blame is the expression of an attitude (or attitudes) that is intended to sting a person because she did something wrong. Examples of such attitudes include resentment, blame, righteous anger, and condemnation. It is important to note, however, that these attitudes can sometimes be expressed without the intention to sting a person because she did something wrong. Scanlon, for example, argues that you can blame

someone by withdrawing a measure of trust or goodwill without intending to hurt the person at all.¹⁸⁵ It seems to me that you can also blame someone by pointing out that their character is flawed in various respects without intending to hurt that person. I call these kinds of blame “benign blame” and I think benign blame plays an important role for many in their lives. I suspect that purported “benign blame” is sometimes abrasive blame in sheep’s clothing, but I trust that it is often sincere (i.e. it honestly involves no intention to sting the wrongdoer). The important point here is that blaming (or resenting) someone is not necessarily abrasively blaming someone. Abrasively blaming someone necessarily involves the intention to sting a person because she did something wrong.

Abrasive blame is distinct from pointing out that a person did something wrong or “calling her out.” Pointing out that a person did something wrong generally involves intending that the person fully recognize and understand that she did something wrong. If a person comes to fully recognize and understand that she did something wrong, she will generally suffer pangs of conscience and guilt. So pointing out that a person did something wrong often involves intending that the person suffer in the recognition of her wrong. Abrasive blame is different; it involves an intention to sting a person above and beyond the pangs of conscience that come with recognition of a wrong. Abrasive blame can (although it need not) continue even after a person realizes that she has done wrong and feels remorse for it. This is because abrasive blame involves an intention to sting from the outside, not (just) to shine light on a truth that gnaws from the inside. For example, suppose I confront a person and calmly tell her that she told a lie. I intend that

¹⁸⁵ See *Moral Dimensions: Permissibility, Meaning, Blame*. Susan Wolf calls this kind of blame “wimpy blame,” (“Blame, Italian Style”).

she recognize her wrong action and thereby feel a pang of conscience, but I do not intend to make her feel any additional hurt. If and when she acknowledges and understands her wrong action I will no longer bring it to her attention. In this case, I do not abrasively blame her. But suppose instead that I call her a “rotten two-faced liar who has got a different story for every pair of eyes.” I hope the words will hurt her and I am prepared to use them regardless of whether she has recognized her wrong action. In this case, I abrasively blame her. I intend to sting her over and above any pangs of conscience she might feel in the privacy of her own heart. I make my remarks not because (or not only because) I want her to recognize a painful truth, but because I want to hurt her for having done something wrong.

An example may help. Consider the following dialogue taken from a scene in the movie *Good Will Hunting*. Will is incredibly intelligent and has had no formal education after high school. He and his friends go to a bar and have the misfortune of meeting a man named Clark who is a first-year graduate student at Harvard. Clark insults one of Will’s friends and Will steps in to defend his friend by outsmarting Clark. Clark tries to shut down Will’s defense by beginning to quote, without reference, a passage from a book on the American colonies prior to the revolution. Will interrupts with the full quote and continues as follows:

WILL: You got that from Vickers. *Work in Essex County*, page 98, right? Yeah, I read that too. Were you going to plagiarize the whole thing for us? Do you have any thoughts of your own on this matter? Or is that your thing, you come into a bar, you read some obscure passage, and then pretend, you pawn it off as your own, as your own idea just to impress some girl and embarrass my friend?

[Thus far we have statements that are arguably intended to make Clark recognize that he has done something wrong. We do not yet have a clear case of abrasive blame. The following remarks, however, are meant to do more than make Clark recognize that he has done something wrong. They are meant to sting him for it.]

You see, the sad thing about a guy like you is that in fifty years, you're going to start doing some thinking on your own and you're going to come up with the fact that there are two certainties in life. One: don't do that. And two: you dropped 150 grand on a f#\$% education you could have gotten for a dollar fifty in late charges at the public library.

CLARK: Yeah, but I will have the degree, and you'll be serving my kids fries at a drive thru on our way to a skiing trip." [This insult is meant to sting, but it is not abrasive blame because it is not meant to sting Will for having done something wrong. Unless, of course, Clark judged that Will had done something morally wrong and uttered this insult to sting him for it. But I highly doubt that this was the case.]

WILL: "That may be, but at least I won't be unoriginal. But I mean, if you have a problem with that, I mean, we could just step outside – we could figure it out."

[This remark is not an expression of abrasive blame. Instead, it is a threat to sting Clark with something other than an attitude (likely a sucker punch or a fist in the face).]

A few more clarifications about abrasive blame are in order before analyzing whether abrasive blame is permissible. Privately thinking nasty thoughts about a

wrongdoer does not count as abrasive blame. This kind of thinking does not generally involve an intention to sting the wrongdoer. Of course there is an exception. If a person (let us call her “the passive-aggressive one”) intends for the wrongdoer to read her mind and be hurt by its mean and nasty contents, then her thinking nasty thoughts does count as abrasive blame. Similarly, saying nasty things about a wrongdoer while talking to a confidante does not count as abrasive blame – unless the person (let us call her “the back stabber”) intends for the wrongdoer to hear and be hurt by her comments through the grapevine. If Bob Dylan were privately singing “Masters of War” instead of recording it for a popular audience that included his target, then the lyrics quoted at the start of the paper would not be an example of abrasive blame. The point is that abrasive blame is the expression of an attitude that is intended to sting the wrongdoer – without mindreading and grapevines (or other sneaky tactics) it has to involve actively communicating something to the wrongdoer.

It is important to clarify the relationship between abrasive blame and criminal punishment. I give a detailed analysis of this relationship in the next chapter, so I limit myself to brief remarks here. Abrasive blame is not synonymous with punishment, nor is it necessarily a part of punishment. Abrasive blame is, however, often a part of punishment. In modern political communities, criminal punishment often involves the expression of attitudes that are intended to sting the offender above and beyond her own pangs of conscience because she did something wrong. But criminal punishment in the modern world also characteristically involves deprivation of liberty. Criminal punishment used to (and often still does) involve torture, death, and other harms that go well beyond the sting of blame.

Someone who expresses abrasive blame intends for the expression of her attitudes to hurt a person because that person did something wrong. Although the abrasive blamer must intend to hurt the wrongdoer, she need not have a calculated or pre-meditated plan to hurt the wrongdoer. Many instances of abrasive blame are not pre-meditated plans to hurt wrongdoers, but are relatively spontaneous or impulsive responses to wrong action.

¹⁸⁶ It is tempting to see such spontaneous responses to wrong action in a benign light. People often respond to wrong action with anger that does not appear to be meant to burn, sting, or hurt anyone in particular. In such cases, harm done to the wrongdoer appears to be an unintended consequence or side effect of the blamer's general anger. Under this benign light, many impulsive acts of blame appear to have thorns that accidentally pierce, but do not intentionally sting. Such acts of blame are not pleasant, but they are not abrasive.

We should beware, however, of the temptation to shine too much benign light on the subject. Sometimes blame pierces in a way that is accidental, sometimes blame has a pre-meditated sting, but many acts of blame lie somewhere in between these extremes. Abrasive blame is like many other human actions; it generally flows from general habits over which we have long-term control.¹⁸⁷ Most people have a general practice, pattern, or habit of responding to wrongs in a relatively consistent fashion. People are generally aware (or can easily become aware) of how abrasive they are in responding to wrongs. People are also capable of changing their practices and habits of blaming behavior. If a

¹⁸⁶ I am indebted to Benjamin Bagley and Susan Wolf for respectively bringing this point to my attention.

¹⁸⁷ I am drawing on a somewhat Aristotelian understanding of human behavior. For criticism of this point, see Doris, *Lack of Character*.

person is aware that her habit of responding to wrong action hurts those whom she is responding to and if she chooses not to change that habit, then her stings of blame are not unintended consequences or side effects. Her blame is abrasive blame that is meant to sting people, not spontaneous blame that just accidentally happens to pierce people.

An example might help illustrate this point. Suppose I have a strong, but not pathological cursing habit and suppose I have a child. I know that cursing in front of the child is harmful for her development, but I do nothing to curb my habit. If my friend were to hear me speak long strings of curses in front of my child, I am certain that she would express disapproval and criticize my behavior. It would be ridiculous for me to protest her disapproval on grounds that I do not *intend* to sully my child's vocabulary. Although creating a little potty mouth is not a cold-blooded, pre-meditated project of mine, it is also not an unintended consequence of my venting frustration with the world. It is intentional in the broad sense of being under my control. It is the result of a habit that I know is bad for my child and that I have the power to change. My friend would rightly criticize me for not changing my habit; it is a bad habit and I should get rid of it.

Let us get back to abrasive blame. I grant that blame is often a spontaneous or impulsive response to wrong action, rather than a calculated sting. This does not mean, however, that spontaneous blame cannot inflict intentional harm. Blame, like cursing, may be impulsive in a moment, but it is intentional in the broad sense of being under our control. Keeping in mind a few exceptions, blame usually flows from a general habit that people can be made aware of and that people can change. If this habit is a bad habit, if it results not just in hurting people but in *impermissibly* hurting people, then people should change this habit and suggest that others do likewise. The prevalence and acceptance of

this habit in our society suggests that people generally do not see anything impermissible about expressing attitudes that hurt people because they have done something wrong.

This point takes us back to the guiding question of this chapter; how can the sting of blame be morally justifiable in the light of a standing injunction against hurting people?

Before addressing this question, I want to briefly contrast it to other questions that philosophers generally ask about blame. Philosophical focus on blame has a tendency to center around freedom and responsibility. What kind of freedom and/or responsibility do attitudes like blame and praise presuppose? Do most people have the presupposed kind of freedom and/or responsibility? Do children, addicts, and psychopaths? These questions are interesting and engrossing, so much so that it is tempting to take other features and questions about blame for granted. It is tempting to assume that – if a person freely and responsibly acts wrongly (and this may be a big “if”) – then it is permissible to blame her. But we should question this assumption. Blame is often intended to hurt people; abrasive blame is meant to sting. This abrasive aspect of blame is at odds with a standing moral injunction against intentionally hurting people. How (if at all) is the sting of blame permissible in light of this injunction? This question, I argue, is worth asking. Like most questions worth asking, it is not easily answered.

Section II: Is Abrasive Blame Permissible?

One might think an easy answer (or simply dismissal of the question) is readily available. Although there is an injunction against hurting people, it is a general, standing injunction that is not without exceptions. The commission of a wrong action is simply another exception to the standing injunction against harm. Although it is usually impermissible to hurt people, it is permissible to hurt them with the sting of blame when

and because they have done something wrong. In committing a wrong action, a person makes herself liable to (at least one kind of) harm – namely, the sting of blame.

I reject this purportedly easy answer. I argue that the commission of a wrong action does not, by itself, make it permissible to abrasively blame the person who committed the action. My argument does not entail that abrasive blame is usually impermissible. I argue that the commission of a wrong coupled with the presence of a certain kind of personal relationship (or community) can make it permissible for one party of the relationship to blame another. To ease exposition of the argument, I give the basic structure in what follows. In the remaining sections and the next chapter, I defend the premises in this argument.

The Basic Argument

1. An act of abrasive blame cannot be justified *tout court*; it cannot be justified on the basis of reasons that one can expect people to accept insofar as they are reasonable. I say why this is the case in **Supporting Argument 1**. The relative weights of reasons for and against abrasive blame are evenly balanced. A reasonable person who conscientiously considers all relevant reasons can decide that abrasive blame is either justified or not.¹⁸⁸
2. If an act cannot be justified on grounds that one can expect people to accept insofar as they are reasonable, then all other things equal, the act is morally impermissible. This claim is an assumption. Scanlon and other moral

¹⁸⁸ I am assuming, following Nagel, that “there is substantial middle ground between what it is unreasonable to believe [or in my case, accept] and what it is unreasonable not to believe [or in my case, reject],” (*Equality and Partiality*, 161).

philosophers have defended it in spirit,¹⁸⁹ if not in specific formulation.¹⁹⁰ It is beyond the scope of my argument to further defend this assumption.

3. So an act of abrasive blame, all other things equal, is morally impermissible.

But other things are not equal.

4. Acts of abrasive blame are often embedded within the context of personal relationships. Parties of these relationships often jointly accept norms that license them to express a reasonable degree of abrasive blame to one another. When this is the case, parties can expect each other to accept a justification of abrasive blame insofar as they are reasonable people *and* reasonable parties of the relationship (or community). I say how this works in **Supporting Argument 2**.
5. Acts of abrasive blame can also be justified within the context of larger communities. I say how this works with respect to political communities in the next chapter, **The Political Sting**.
6. Within certain personal relationships and communities, parties can permissibly engage in acts of abrasive blame toward one another.

¹⁸⁹ See Scanlon, *What We Owe to Each Other*. I depart from Scanlon because I do not agree that the following two statements are equivalent: 1) an act is wrong if any principle that permitted it could reasonably be rejected and 2) an act is wrong if it would be prohibited by any principle that people could not reasonably reject. I believe my departure from Scanlon here stems from his apparent reluctance to affirm cases of persistent reasonable disagreement (*What We Owe to Each Other*, 196). Unlike Scanlon, I readily affirm that there are cases in which a principle *permitting* some action X could reasonably be rejected and a principle *prohibiting* the same action X could reasonably be rejected. This departure leads me away from using Scanlon's strict formulation of impermissibility in terms of reasonable rejection. Instead, I begin my argument with an assumption that Scanlon's theory defends in spirit, but not in detail.

¹⁹⁰ *What We Owe to Each Other*, 4.

Before providing the supporting arguments, I want to emphasize two features of the basic argument to avoid confusion. First, philosophers often consider abrasive blame in the same bundle as praise, benign blame, and other reactive attitudes. The permissibility of commendation and condemnation, laud and reproach, resentment and guilt are sometimes thought to stand or fall together.¹⁹¹ This is likely due to the fact that philosophers tend to focus on whether or not these attitudes presuppose a kind of freedom or responsibility that people may lack. As previously mentioned, my focus is different; it is not on the freedom/responsibility that blame (or praise) may presuppose, but on the hurtfulness that blame can have. My argument is silent on praise, benign blame, and any other response to action that does not involve the intentional infliction of a hurt. These responses lack the sting of abrasive blame and it is disagreement about whether “the sting” is justified that gets my argument off the ground in the first premise.

Second, my argument is about the moral permissibility of engaging in abrasive blame, not about a political right to engage in abrasive blame. A just society should protect a right to freedom of expression. It should not coerce people or punish them for abrasively blaming one another just like it should not punish or use force to prevent people from expressing insults, swear words, or offensive images. A just society ought to defend a person in her claim to say things like, “you dirty rotten bastard!” But of course, there is a difference between defending a person’s right to express something and arguing that it is morally permissible for a person to express a particular thing. I defend a

¹⁹¹ Galen Strawson seems to suggest as much in “The Impossibility of Moral Responsibility.” All these attitudes seem to build in a “desert-basis” which is compromised by psychological determinism.

person's right to express abrasive blame, but I have argued that she should not exercise it unless it is in the context of a certain personal relationship or community.

Section III: Reasonable Disagreement on Abrasive Blame (Supporting Argument 1)

In this section I argue that an act of abrasive blame cannot be justified on the basis of reasons that one can expect people to accept simply insofar as they are reasonable. I show that a reasonable person who is conscientiously considering reasons for and against abrasive blame can decide that abrasive blame is either justified or not.

I begin with what I have referred to as a standing moral injunction against acting on an intention to hurt a person. This standing injunction is not a categorical prohibition against acting on intentions to hurt people. Such a prohibition would be unreasonable; it would prohibit dentists from performing root canals (and routine cleanings for that matter) and would-be murder victims from fending off their attackers with pepper spray or stern kicks to strategic locations. The standing injunction is rather to be interpreted in the following fashion: there is (at least) presumptive moral reason to refrain from intentionally hurting a person. Other reasons, such as preventing future pain or bringing about various moral goods, might defeat the presumptive reason to refrain from intentionally hurting a person. Some conditions, such as consent from the individual who is to be hurt, might undermine the reason to refrain from intentionally hurting a person. Giving a full analysis of all the potential defeating reasons and undermining conditions is well beyond my scope. My focus is on the act of abrasive blame – when a person engages in abrasive blame, she intentionally hurts a person. My question is: are there reasons to engage in abrasive blame that defeat or undermine the reason to refrain from intentionally hurting a person? I argue that this question has a range of reasonable

answers. In what follows, I describe several reasonable answers (the list is not meant to be exhaustive).

The Qualified Retributivist

Although people have a general reason to refrain from intentionally hurting a person, they do not necessarily have a reason to refrain from expressing attitudes that are intended to inflict pain that a person *deserves*. As long as a wrongdoer had an adequate opportunity to avoid doing wrong, she morally deserves to experience (at least) some pain. She does not deserve to experience more pain than is proportional to her wrong action, but the pain of abrasive blame rarely exceeds this amount. Just because a wrongdoer deserves to feel pain, however, does not mean that an individual is morally licensed to inflict this pain in any fashion. For example, if a criminal were to deserve imprisonment, this fact would not license any individual to lock him up in their closet. An individual is licensed to inflict morally deserved pain on others only if she has the right kind of social authority or standing to inflict the deserved pain. Individuals generally lack the requisite authority or standing to inflict deserved pain in the form of physical injuries, imprisonment, or monetary fines.

Individuals do, however, generally have the requisite authority or standing to inflict deserved pain in the form of expressing hurtful attitudes. Recall the dialogue from *Good Will Hunting* quoted above. Intuitively, Will does not need any special standing to sting Clark with words and facial expressions. Abrasive blame does not intend to inflict just any pain in any fashion; it intends to inflict deserved pain through the expression of hurtful attitudes. The armor of morality need not shield our skin from every single prick. It need not shield us from pricks that we deserve and that come in the form of words and

facial expressions. Although there is generally moral reason to refrain from hurting people, there is generally not *prima facie* moral reason to refrain from abrasive blame.

This argument relies on the claim that wrongdoers morally *deserve* to feel at least some pain for their wrong actions as long as they had an adequate opportunity to avoid committing them (this is why I label a holder of the view a “qualified retributivist”).¹⁹² Although it is reasonable to reject this claim, it also seems reasonable to accept it. It is supported by a set of strong intuitions; I present two such intuitions. Consider a world in which wrongdoers experience some pain that is less than or equal to the amount proportional to their wrong. Compare it with a world in which wrongdoers receive no pain. Most people have the intuition that the former world is no worse than (and may even be preferable to) the latter world. Although there is less pain in the latter world, this fact does not seem make that world preferable from a moral point of view – presumably because the pain that is missing is pain that wrongdoers morally deserve. Consider another case: if two innocent people and one murderer are on a lifeboat that will sink unless one of them is thrown off, then intuitively the murderer is the person who should be thrown off. Intuitively we take ourselves to have less reason to avoid inflicting pain on guilty parties compared to innocent parties – presumably because guilty parties deserve (at least some) pain.

One might object to the use of these intuitions to support the qualified retributivist’s view. Perhaps these intuitions are parasitic on a set of commonly felt retributive emotions that are, like racial prejudices, pathological in their very nature. This objection has its roots in Nietzsche’s powerful critique of the retributive emotions. He

¹⁹² Michael Moore, among others, defends a view along these lines in *Placing Blame*.

argues that these emotions are rooted in *ressentiment*. Those who feel *ressentiment* are weak and jealous of the strong. They make up the idea that people deserve to feel pain in order to feel good about themselves when they inflict pain on the “deserving.” A number of philosophers have adequately responded to this critique by arguing that retributive judgments are not inevitably motivated by *ressentiment* or other pathological emotions. For example, Michael Moore argues that retributive judgments can flow from the same non-pathological, emotional root as personal guilt.¹⁹³ Jean Hampton has reasonably argued that retributive judgments are a way of standing up for morality.¹⁹⁴ Although these philosophers do not have abrasive blame in mind, their arguments can be adapted to justify abrasive blame along qualified retributivist lines.

The Hard Incompatibilist

The hard incompatibilist (at least the kind I am describing for the purpose of my argument here) is a variant on the qualified retributivist. The hard incompatibilist holds that it is permissible to abrasively blame a wrongdoer insofar as she deserves it but she deserves it only if she had an adequate opportunity to avoid doing wrong.¹⁹⁵ Hard incompatibilists (among others) argue that all wrongdoers lack an adequate opportunity to avoid doing wrong. Others argue that we have strong reason to doubt and cannot be reasonably sure that wrongdoers have adequate opportunities to avoid doing wrong. The longevity and tenacity, but most importantly the quality, of the free will debate indicates

¹⁹³ Michael Moore, “The Moral Worth of Retribution.”

¹⁹⁴ Hampton, *Forgiveness and Mercy*.

¹⁹⁵ Scanlon maintains that an adequate opportunity to avoid doing wrong may not be sufficient to justify abrasive blame (especially if the sting is particularly severe and prolonged), but it is at least necessary to justify abrasive blame (*Moral Dimensions*, 187).

that such positions are reasonable; they are internally consistent and responsive to objections (which is not to say that a reasonable person must accept these positions; after all it is a *debate*). According to this position, desert could justify abrasive blame but we cannot actually appeal to claims of desert in the actual world. According to this position, abrasive blame may still be justified according to some other argument, but not along the lines that the qualified retributivist presents.

The Qualified “Pacifist”

In contrast to the argument above, the following argument suggests that individual acts of abrasive blame are unjustified. The term “pacifist” generally refers to a person categorically opposed to war, but I use the term here to refer to a person who takes there to be a very strong moral reason to refrain from increasing pain and suffering of people in the world.¹⁹⁶ The pacifist maintains that abrasive blame increases the amount of pain and suffering people experience. She maintains that there are no *prima facie* reasons in favor of abrasive blame that are strong enough to defeat the very strong reason to refrain from increasing pain.

I cannot prove the pacifist’s claim that there are *no* reasons in favor of abrasive blame that can outweigh the strong reason to refrain from increasing pain. I do, however, defend the pacifist in the following fashion. I consider the reasons in favor of abrasive blame that are the best candidates for outweighing the reason to refrain from increasing pain. I argue that the pacifist can reject these reasons.

¹⁹⁶ This opposition to unnecessary pain is showcased in utilitarianism, particularly negative utilitarianism, but it is compatible with other moral theories.

First, one might think that the pain caused by abrasive blame is somehow different from other pains. It is the kind of pain that can be a morally good kind of pain to bring about, or at least not a bad kind. Someone might support this claim, as the retributivist supported her claim, by appealing to intuitions. Compare a world in which wrongdoers receive the pain of abrasive blame that is proportionate to their wrongs to a world in which wrongdoers do not receive the pain of abrasive blame. If one has the intuition that the former world is morally preferable, then one might think that the pain of abrasive blame is a good kind of pain to bring about. Perhaps some people have such an intuition (and the pacifist need not judge them pathological!), but this is certainly not an intuition that one *must* have insofar as one is a reasonable person. The pacifist can reject it and maintain that there is nothing morally special about the pain inflicted by abrasive blame; it is like other pains, and we should not act in ways that increase it.

Second, one might object to the pacifist by arguing that engaging in abrasive blame promotes, defends, or instantiates a moral ideal or value that outweighs the strong reason against inflicting pain. There are two potential values that engaging in abrasive blame may promote or instantiate. First, perhaps engaging in abrasive blame may promote or instantiate the morally valuable self-respect of the victim of wrongdoing. The pacifist can reasonably insist, however, that there are reasonable methods to restore and defend the self-respect of the victim that do not involve intentionally hurting another person. Given the strong reason against inflicting pain, the victim ought to use the non-abrasive methods of defending and restoring self-respect. On a rare occasion, these non-

abrasive methods may not be available, but that is certainly an exception, not the rule.¹⁹⁷ In almost all cases, a victim can stand up for herself without pushing someone else over.¹⁹⁸

Perhaps, however, abrasive blame helps defend, promote, or instantiate some moral value other than self-respect. One might argue that abrasive blame instantiates a way of caring about morality itself in the face of immoral actions.¹⁹⁹ But this argument is unconvincing. Surely one can take morality seriously in the face of an immoral action by expressing that the act in question was wrong and by encouraging the wrongdoer not to repeat his mistake. To refrain from abrasively blaming a wrongdoer is not to condone wrongdoing. There are other non-abrasive ways to hold people accountable to moral standards. The qualified pacifist need not make the strong move of denying that abrasive blame is a way of caring about morality. She simply thinks there are alternative ways to care about morality that introduce less pain into the world. Given her firm moral commitment against increasing pain, she holds that abrasive blame is not justified in light of the presence of these alternatives.

The pacifist is not yet “in the clear” in her argument that abrasive blame is unjustified. One can object to the pacifist’s argument on the following pragmatic grounds. The qualified pacifist claims that abrasive blame hurts people in a way that

¹⁹⁷ Nietzsche’s discussion of the will to punish may explain why people are often inclined to think that restoring self-respect requires hurting wrongdoers (*On the Genealogy of Morality*). I insist, however, that this thinking is misguided; a self-respect worth having is generally not contingent upon hurting others.

¹⁹⁸ Contra Hampton, *Forgiveness and Mercy*.

¹⁹⁹ Jay Wallace, “Dispassionate Opprobrium.” George Sher also defends a view along these lines, see *In Praise of Blame*.

increases pain in the world. This claim may be false. Although abrasive blame is intended (and often succeeds) to inflict pain upon the wrongdoer in the moment, it may decrease overall pain in the long run by deterring the wrongdoer and others from committing future wrong acts that cause pain. Abrasive blame may also decrease the pain felt by the victim of the wrongdoing. It may help the victim vent her anger, or provide an easier, less painful way to restore or maintain her self-respect.

The pacifist can respond to the objection in a reasonable fashion. She can grant that an act of abrasive blame does not increase pain (and may even decrease it) when it is compared to *condoning* wrong action. But abrasive blame may very well increase pain when it is compared to other non-abrasive responses to wrongdoing – such as engaging in benign blame, providing criticism that is intended to educate rather than sting, granting forgiveness, and giving support and encouragement to do better. Empirical evidence is inconclusive on this point, but a substantial amount of evidence suggests that forgiveness (which is devoid of abrasive blame) is generally more effective than resentment (which usually involves abrasive blame) at both deterring future offenses and helping victims heal and recover. Given the absence of definitive evidence to the contrary, it is reasonable for the pacifist to maintain that abrasive blame increases pain compared to other methods of responding to wrongdoing. Given that such methods are generally available, it is reasonable for the pacifist to conclude that abrasive blame is generally unjustified.

The Tough-Love “Pacifist”

The above argument has revealed that it is also reasonable for a person to maintain the opposite stance. The empirical evidence is inconclusive on the relative

amount of pain that abrasive and non-abrasive responses to wrongdoing cause in the long run. A person, let us call her the “tough-love pacifist,” can review the evidence and reasonably conclude that abrasive blame does not increase pain when compared to non-abrasive methods of responding to wrong action. The tough-love pacifist, like other pacifists, recognizes a very strong moral reason against increasing pain in the world. She maintains, however, that abrasive blame does not increase pain in the world; it is (equally or) more effective than the relevant alternatives at minimizing pain in the world. As long as the tough-love pacifist sees no other grounds on which abrasive blame may be wrong, she will take it as a justified response to wrongdoing (perhaps even a required response in order to minimize pain in the world).²⁰⁰

Although a pacifist described above disagrees with this argument, I do not think she can prove that it is an unreasonable argument. She can reasonably reject principles that permit abrasive blame, but she does not have sufficient grounds to preclude others from reasonably rejecting principles that prohibit abrasive blame.

An Objection

I have presented four positions that are internally consistent and able to respond to objections, but I have not shown that each is reasonable according to an explicitly shared standard of reasonableness. Rather than providing such a standard, I identify the types

²⁰⁰ My argument in “The Political Sting” is similar to this argument. But there is an important difference. The tough-love pacifist reasonably claims that engaging in abrasive blame minimizes pain compared to the alternatives. She does not have sufficient evidence, however, to reasonably refute those who disagree. Although she has *some* reason to hold that abrasive blame minimizes pain, she lacks *conclusive* reason to hold that abrasive blame minimizes pain. In “The Political Sting,” I argue that there is conclusive reason to hold not that acts of abrasive blame minimize pain, but that maintaining social rules about abrasive blame in a political community minimizes pain.

of disagreement that separate the various positions from one another. Then, drawing on the tradition of political liberalism, I argue that respect for persons requires us to respect these types of disagreement. From where we stand, we cannot reasonably expect all persons conscientiously considering the relevant considerations to arrive at the same answer on whether (and how) the commission of a wrongdoing justifies abrasive blame in the absence of special relationships or communities.

I see three different nodes of disagreement in arguments about whether and how abrasive blame is justified and I will consider each node in turn (see Figure 1 for my characterization of the debate and its three nodes of disagreement). The first node of disagreement flows from a difference in basic value commitments. The difference can be identified in different answers to the following question: would committing a wrong in the presence of an adequate opportunity to avoid doing so make it the case that a person deserves to be stung with abrasive blame? By “deserve to be stung” I mean the following: a person deserves to be stung if the fact that she committed a wrong action makes it the case that the sting is not a morally bad pain for her to suffer. When a person deserves to be stung, the moral presumption against inflicting pain in the form of the sting is undermined because such pain is morally innocuous (and may even be morally good). The qualified retributivist and the hard incompatibilist think that the commission of wrong in the presence of an adequate opportunity to avoid doing wrong does make it the case that a person deserves to be stung. The qualified pacifist and the tough-love pacifist hold that it does not. The difference in their answers flows from a general difference in their comprehensive moral (and/or religious) views, particularly in their conception of value. The retributivist and the hard incompatibilist see value in a

wrongdoer suffering above and beyond the pangs of conscience that come with a recognition of having done wrong. The pacifists see no value in this suffering; like other kinds of suffering, it is an evil to be avoided.

The difference between these views often (but not necessarily) gets played out in the various players' views about hell as understood in the mainstream Judeo-Christian tradition.²⁰¹ The pacifist rejects the possibility that a morally perfect being could send persons to hell. Her rejection is rooted in a commitment to seeing all suffering beyond internal pangs of conscience as evil. The qualified retributivist might remain open to the possibility of hell. Although she would be concerned that an eternity of suffering is disproportionate to most offenses, she does not share the pacifist's commitment to seeing any suffering beyond pangs of conscience as evil. The qualified retributivist thinks wrongdoers deserve suffering above and beyond pangs of conscience; she takes such suffering to be innocuous or even valuable. Although it may not be pretty (albeit poetic) for a being to send people to hell (at least for some period of time), it is also not wrong or evil.

The pacifist and the qualified retributivist do not share a common ground from which one of their views is decidedly better. Neither is making a mistake in drawing inferences from shared premises. Instead, the split between the pacifist and the qualified retributivist flows from a deep difference in value. This kind of difference should be respected, not forcibly erased. If a qualified retributivist respects that the pacifist has the moral power to autonomously develop her own conception of the good and the pacifist has indeed conscientiously and autonomously come to her view, then the retributivist

²⁰¹ Galen Strawson makes a similar move in "The Impossibility of Moral Responsibility."

must respect the pacifist's view. Compelling her to change it would involve compelling her to change her more comprehensive understanding of value in which her view is situated. The tradition of liberalism – rooted in respecting the individual's power to develop her own conception of the good – stands strongly against such compulsion.

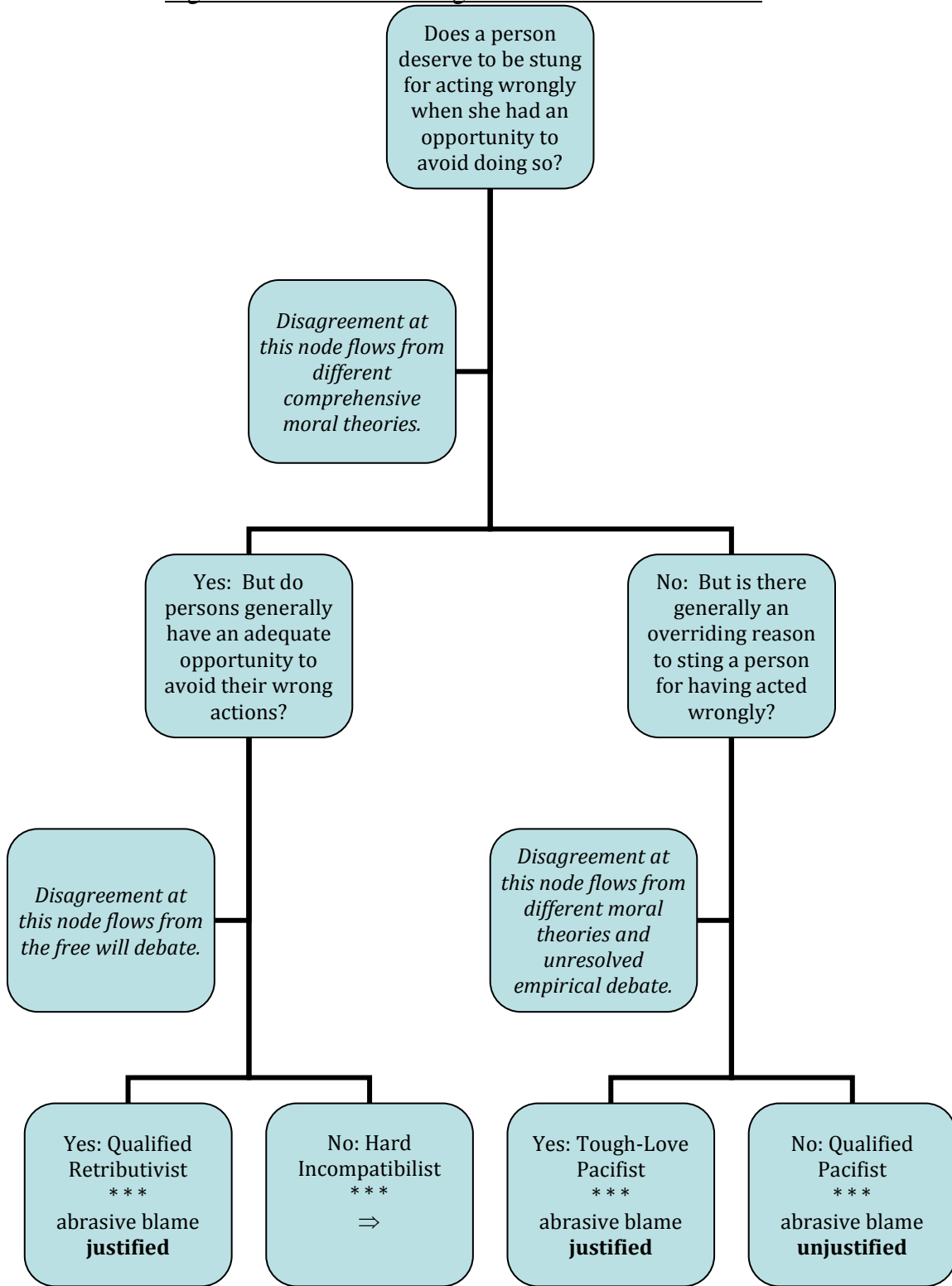
This argument, however, does not necessarily entail that there is reasonable disagreement about whether abrasive blame is justified by the commission of a wrong. I have thus far only shown that there is reasonable disagreement on whether the suffering inflicted by abrasive blame is deserved when a person commits a wrong in the presence of available opportunities to avoid committing the wrong. If all those who think people deserve a sting for acting wrongly in the presence of adequate opportunities to avoid doing wrong *deny (or must deny) that people have such opportunities*, then regardless of the value disagreement with the pacifist, there may be no disagreement about whether abrasive blame is justified in actual cases. (Figure 1 helps illustrate this point.) But the antecedent here is false. The hard incompatibilist's arguments are not strong enough to show that a reasonable person must deny that people have adequate opportunities to avoid doing what they do (in our case, doing wrong actions). It is rational and reasonable to maintain that such opportunities exist for human agents and that we are in a position to know whether such opportunities are available to an agent with respect to a given action. I do not deny that the existence and knowledge of these opportunities is subject to debate, but I do claim that such debates are *genuine debates*. As the history and quality of the free will debate indicates, the hard incompatibilist lacks sufficient evidence to decisively shut down the arguments of what I have called the qualified retributivist.

If one takes the qualified retributivist's view, then there is no moral presumption against abrasive blame. All other things equal, abrasive blame is justified in response to a wrong. But one can respectably depart from the retributivist's value judgment about deserved suffering or her judgment that people have adequate opportunities to avoid committing the wrong actions they commit. If so, then the presumption against abrasive blame remains. Even if the presumption stands, however, abrasive blame may still be justified in response to a wrong. When a person commits a wrong action, there may be general reasons to abrasively blame her that outweigh the presumption against abrasive blame. For example, stinging the person with abrasive blame may help the person flourish in the long run. One might think that being stung with abrasive blame is like having a root canal; it hurts but the presumption against the hurt is outweighed because it brings about long-term benefits. Perhaps abrasive blame helps the offender engage in moral reform that improves her life in the long run. In addition, abrasively blaming the offender may help deter her (or others) from committing future offenses. If abrasive blame has such benefits, they might be worth the cost of inflicting the sting. But again, this claim is subject to reasonable disagreement. The disagreement at this node is both empirical and moral. Data is inconclusive on whether the sting of blame deters future offenses or helps improve offenders' lives. Moreover, interpretation of the data is complicated by the fact that we disagree on what counts as an improvement to a person's life. Most importantly, there is reasonable disagreement on whether future benefits (to the offender or others) can justify the infliction of present harm in the absence of consent or special relationships. As it did in the first node, the disagreement here is situated in broader disagreement over comprehensive moral theories: most utilitarians would agree

that present harm is justified for future benefits, but needless to say there are other reasonable moral theories that reject such a view. Once again, there is disagreement here that we must respect rather than erase.

In sum, there is disagreement that we should respect on whether abrasive blame is justified in response to wrong action. Some maintain that it is justified because deserved, others reject that it is deserved. Of those who think abrasive blame is not deserved, some think it is justified by countervailing benefits and others reject that benefits cannot overcome the presumption against infliction of undeserved harm. Each of these positions is internally consistent and responsive to objections. Moreover, the type of disagreement that holds among these positions is the type we should respect.

Figure 1: Reasonable Disagreement on Abrasive Blame



At this point, I hope to have shown that reasons do not conclusively count for or against abrasive blame in a way that we can force a reasonable person to accept. A reasonable person who conscientiously considers all the relevant reasons can decide that abrasive blame is morally permissible or not. If she holds one of the positions described above, we must respect this fact and not demand that she change her position. The fact that a person is *reasonable* is not a sufficient ground on which to expect that she accepts a justification of abrasive blame. Following the basic argument sketched above, I conclude that abrasively blaming her would, other things equal, be impermissible.

I offer an example to help illustrate this conclusion. Suppose Bob does something wrong to Jane. Jane considers stinging him with abrasive blame. Perhaps if she were in Bob's shoes, she would think abrasive blame is justified. But she is not in Bob's shoes; Bob is. Bob could make a good case that abrasively blaming him is wrong (perhaps he holds the pacifist position I described above) and Jane could not discredit such an argument. All she could do is maintain that she "respectfully disagrees." Abrasive blame should not follow from reasonable disagreement. Abrasively blaming Bob would *hurt him* under the pretext of reasons that he cannot be required to accept.²⁰² That would not be "respectful disagreement." It would be disagreement that turns out to hurt Bob and not Jane; an asymmetrically hurtful sort of "respect" indeed.

In this example, Bob and Jane are very abstract characters. As the case is described, all that they can reasonably expect from one another is that they are reasonable

²⁰² If Jane could make a case that refraining from abrasive blame would hurt her, then she might have a leg to stand on. She might argue that abrasively blaming Bob is the only way for her to recover her self-respect after having been wronged by him. As previously discussed, however, reasonable methods to restore self-respect that do not involve intentionally hurting another human being are usually available.

people. When all other things are equal and when no other reasonable expectations are in place, then abrasive blame is impermissible between them. But the world is much more complicated. If Jane and Bob are best friends, for example, then they reasonably expect a great deal more from one another than bare reasonableness. As I argue in the next section, they may have developed their own “ways of doing things” that they reasonably expect one another (but not other people) to accept. Often these “ways of doing things” permit abrasive blame.

Section IV: Abrasive blame in personal relationships (Supporting Argument 2)

In this section, I argue that abrasive blame can be permissible within certain personal relationships – namely those that maintain a shared acceptance of a norm that licenses reasonable degrees of abrasive blame. My aim here is to explain (in a bunking, not de-bunking fashion) the common intuition that it seems “obviously right” for us to abrasively blame those with whom we have a personal relationship for the sole reason that those people have done something wrong. Although it often seems right to cite the bare commission of a wrong as a justification for abrasive blame in personal relationships, we should be aware that the justification is conditional upon those relationships. Therefore we should not (at least not without a great deal of caution) import intuitions about abrasive blame in personal relationships into analysis of abrasive blame toward those with whom we have no personal relationship. This sets up my analysis in Chapters 4 and 5 of the political community’s institutional expression of abrasive blame toward offenders (as I explain later, there is no personal relationship in this context).

What I call a *personal relationship* holds when persons²⁰³ have a shared²⁰⁴ acceptance of norms governing their interactions that are somehow *different* from the set of norms a person is required to accept insofar as she is reasonable.²⁰⁵ To explain what this means, I need to say what I mean by “norms.”²⁰⁶ Norms are standards that are meant to guide behavior. They provide a basis for assessing behavior as correct or incorrect. When people accept a norm, they take the norm to guide their action. They judge that they ought to act in ways that the norm deems correct and refrain from acting in ways that the norm deems incorrect. They accept responsibility for acting in accord with the norm and they assess one another’s actions on the basis of the norm. There are some norms, like basic moral principles, that people are expected to accept insofar as they are reasonable. For example, we are expected to accept a norm prohibiting murder insofar as we are reasonable. There are other norms, however, that people can reasonably accept, but they can also reasonably reject. For example, many norms of fashion can be

²⁰³ An individual person may have a personal relationship with herself if she accepts norms other than the ones she is required to accept for governing her own behavior.

²⁰⁴ By *shared* acceptance, I mean that they do not independently accept the same norm as a matter of personal policy. They have a sense that the norm is shared, that it is *their* policy on how to treat one another. I sometimes use the terms “joint acceptance” or “jointly accept” to refer to a shared acceptance. I owe this point to Postema, “Custom in International Law.”

²⁰⁵ I am not arguing that a jointly accepted, differentiated set of norms is *the* constitutive feature of what we paradigmatically call a personal relationship. A good answer to, “what makes up your relationship with your mother?” is not (at least not only) “we govern our interactions with slightly different norms than we are required to accept.” Rather, the presence of a differentiated system of norms picks out the class of what I call personal relationships for the purpose of the paper. Carving out this class of personal relationships may not be particularly interesting in and of itself, but I think it helps us gain important insight into the permissibility of abrasive blame.

²⁰⁶ In this discussion, I draw on Postema, “Custom in International Law.”

reasonably accepted but also reasonably rejected. When people accept norms for governing the interactions between or among them that they are not required by reason to accept, then they have what I call a personal relationship.

Sometimes parties of a relationship jointly accept a norm that licenses reasonable degrees of abrasive blame. Call such a norm an abrasive blame norm. Drawing on arguments in Section II, we know that people are not required by reason to accept such a norm. We also know that people are not required by reason to reject such a norm – so long as the norm does not allow unreasonably high amounts of abrasive blame.²⁰⁷ If parties of the relationship together accept an abrasive blame norm, then they take some degree of abrasive blame to and from the other to be correct. The norm they accept guides the expressions of abrasive blame between (or among) them. One would legitimately tell the other that she is “out of line” when she has abrasively blamed in a way that the norm deems incorrect. When one party abrasively blames the other in a way that stays “in line” with their abrasive blame norm, she legitimately expects the other to accept the blame as justified – not because she is reasonable but because she is reasonable *and* a participant in their relationship with its own shared norms. Abrasive blame between the parties is morally appropriate as long as it is done in accord with their shared norm and their shared norm is not unreasonable.

There are two ways people might come to share acceptance of an abrasive blame norm in their personal relationships. First, they might give one another explicit or

²⁰⁷ Some people (perhaps members of Emperor Nero’s inner circle or members of inner city gangs today) might have a shared norm that licenses unreasonably high amounts of abrasive blame. These norms are not, strictly speaking, abrasive blame norms in the way I have defined them. Abrasive blame norms only license reasonable degrees of abrasive blame.

implicit consent in the form of a contract. Although it is certainly possible to do this, it is highly unlikely that people accept an abrasive blame norm in this fashion. Much more commonly, members of a relationship come to share an abrasive blame norm by undertaking a mutual commitment that emerges from and is sustained by practice over time. By interacting with one another and assessing one another's behavior over time, they develop and sustain a shared understanding of what is correct and incorrect within their relationship. It is not a matter of picking up on a pattern of behavior in the relationship and simply repeating it (they are not training each other to be good dogs), but rather picking up on a shared understanding of what counts as the right way to treat one another.²⁰⁸ When this happens, parties of a relationship share an acceptance of an abrasive blame norm.

If a person jointly accepts an abrasive blame norm in one of her relationships, this does not entail that anyone can justify abrasively blaming her according to that norm. It is reasonable, advisable, and indeed likely, for a person to jointly accept an abrasive blame norm for governing her interactions with one person (such as her mother) and not accept any abrasive blame norm for governing her interactions with another person (such as the clerk at the grocery store). Also, a person may jointly accept a very different abrasive blame norm to govern her interactions with one person (such as her employees)

²⁰⁸ A great deal of subtle social and moral psychology would be needed to determine exactly how and when people have developed such shared understandings of what counts as the right way to treat one another. This work is beyond the scope of my dissertation. My ultimate aim is not to pick out when abrasive blame is justified in personal relationships. My focus is on abrasive blame that political communities express through their criminal justice systems. As I explain in the next chapter, in diffuse, pluralistic political communities there is no such shared acceptance of abrasive blame norms. Insofar as abrasive blame is justified in this context, the justification proceeds in an entirely different fashion.

compared to another person (such as her best friend). The abrasive blame norms in our relationships may differ in the severity of abrasive blame they allow for and in which wrongs legitimately trigger the expression of abrasive blame. For example, the abrasive blame norms in my relationships are such that it is permissible for my close friends to abrasively blame me in ways that it is impermissible from my professional acquaintances to abrasively blame me.²⁰⁹

The account I am offering here helps explain a common intuition that very similar expressions of abrasive blame in response to the same wrongdoing can be wrong coming from some people, but morally legitimate coming from others. Consider the following hypothetical scenario. Imagine that you lie to a close friend about something important. Your friend finds out and she confronts you about it, stinging you with blame by scowling and telling you that you are a total piece of scum. Meanwhile, a stranger has also found out about your lie. One day you see this person at a coffee shop and she scowls at you, saying that you are a piece of scum for having lied to your friend. Intuitively, at least to me, your friend's response is legitimate, but the stranger's response is wrong. If the moral permissibility of blame were dependent only upon people being free and responsible agents, then it seems there should not be anything wrong with the stranger's response. I hope to have explained, however, that other conditions need to be satisfied in order for blame of the abrasive kind to be permissible.

Section V: Contrast to Scanlon

²⁰⁹ In Neil Young's words, "There ain't nothing like a friend who can tell you you're just pissin' in the wind," ("Ambulance Blues").

My account echoes the work of several recent philosophers who have convincingly argued that our practices of blame are deeply dependent on personal relationships. Personal relationships obviously affect who and how much we should blame. It is less obvious that our personal relationships determine the nature of blame and the norms that govern it. T.M. Scanlon takes this far-reaching approach in his recent book, *Moral Dimensions: Meaning, Permissibility and Blame*. Although Scanlon's account offers crucial insight, it does not pay sufficient attention to abrasive blame.

According to Scanlon's theory, an agent is blameworthy for an action if the action demonstrates that there is something faulty about her attitudes that impairs the relationships that she has and can have with others. To judge that a person is blameworthy, however, is different from blaming her. Scanlon argues that blame is neither an evaluation of character nor a sanction. Instead, blame is making a judgment of blameworthiness and accordingly making an appropriate kind of change in a relationship. To blame an agent is to modify one's relationship with her in a way that is made appropriate by her impairment of the relationship. "To blame someone is to hold attitudes toward him that differ, in ways that reflect the impairment, from the attitudes required by the relationship one would otherwise have with the person."²¹⁰ One can appropriately modify one's relationship with a blameworthy agent by trusting her less, denying (or wishing her less) goodwill, and withdrawing in other ways from her. Scanlon maintains, however, that intentionally inflicting distress or suffering upon a wrongdoer is not an appropriate modification of one's relationship with her. People may call this "blame," but they should recognize that this response is very different from his

²¹⁰ Scanlon, *Moral Dimensions*, 145.

concept of blame; it is something meant to hurt the blameworthy agent as an individual, *not* a modification of the relationship that holds between you and the blameworthy agent. Scanlon argues that this kind of response is very difficult, if not impossible, to morally justify.

I agree with Scanlon's distinction between intentional infliction of distress or suffering upon a wrongdoer and what he calls "modification of a relationship." And I agree that the former is very difficult to morally justify. But I disagree with Scanlon's decision to set aside this kind of response to wrongdoing as something distinct from blame properly construed.²¹¹ I think this response to wrongdoing is a very important species of blame; it is what I call abrasive blame. It is very common in our human experience and it is almost always recognized as blame. If an analysis of blame does not capture abrasive blame *as* blame, then I think that analysis has missed something important.

Scanlon's account succeeds in explaining and justifying one flavor of blame, but he fails to provide a full account of blame because he leaves out the "nasty" flavor that is undeniably a part of human life. I recognize abrasive blame as blame and so I maintain, contra Scanlon, that at least some blame *makes sense* outside the context of relationships. We can define and understand abrasive blame without appealing to a personal relationship (or even a bare moral relationship). But similar to Scanlon, I argue that blame is *permissible* only within personal relationships, and as I discuss in the next

²¹¹ This is not the only problem with Scanlon's account. I think he is mistaken in arguing that when a person acts in a way that indicates a faulty attitude, she thereby "impairs" her personal relationships with others. I think the person violates norms of her relationship, but I see no reason to think that such violations necessarily impair a relationship. I suspect that this has to do with the difference in how I define a relationship.

chapter, certain communities. In the next chapter, I discuss the permissibility of blame in a political community, but the argument could be adapted to apply to various other communities.

Conclusion

The moral terrain regarding abrasive blame is indeterminate. In this indeterminate terrain, a person ought to bear on the side of not harming others with the sting of blame. She does not have sufficiently strong reasons to undermine (or override) the standing moral injunction against harming others. Abrasive blame is impermissible in this indeterminate terrain. People are not, however, stuck wandering in this indeterminate terrain. When they form relationships with one another, they tend to navigate and together carve their own determinate paths through this terrain. Some relationships carve paths that allow for a good deal of abrasive blame, others carve paths that allow for a more limited amount of abrasive blame, and still others do not allow any abrasive blame. Relationships are a condition (not the only condition) in which abrasive blame can be permissible. In the next chapter, I show that a community, particularly a political community, is also a condition in which abrasive blame can be permissible.

CHAPTER 4: THE POLITICAL STING OF BLAME

“You can spit in my face once or twice and it is nothing. You can take something away from me and I can learn to live without it. But you cannot spit in my face every day for ten thousand days...and say it is nothing.”

~Jack Henry Abbott²¹²

In the previous chapter I identified abrasive blame (the intentional expression of attitudes meant to hurt a person because she did something wrong) and discussed its moral permissibility in the absence of any special social conditions as well as in the context of a personal relationships. In this chapter, I investigate its moral permissibility in the context of a political community. In Section I of this chapter, I explain why an investigation of abrasive blame is relevant to normative analysis of modern criminal justice systems. I argue that particularly powerful and hurtful stings of blame accompany criminal punishment in modern political communities, especially in America. It is easy to take these stings for granted and focus on other morally disturbing aspects of modern criminal punishment. We should not, however, overlook the sting of blame in criminal punishment; we should not “say it is nothing.” Can we say it is morally permissible?

In Section II, I argue that when a political community has decent social rules allowing for abrasive blame, it is *prima facie* morally permissible for the criminal justice system to abrasively blame many offenders. I argue in the next chapter, however, that

²¹² *In the Belly of the Beast*, 37. Abbott committed suicide in prison in 2002 after over 30 years of incarceration.

even in such a context, it is prima facie morally impermissible for the political community to blame some offenders (Abbott happens to be among them).

My approach here bears structural similarities to the approach Rawls takes up with respect to distributive justice.²¹³ Rawls claims that people do not morally deserve high salaries for productive work, and I claim in “The Sting of Blame” that people do not morally deserve abrasive blame for committing wrongs (at least not in a sense we can reasonably agree to and act upon). Rawls argues that under just background institutions, workers can be legitimately entitled to salaries even though they are not morally deserving of them. I argue in Section II of this chapter that when just *social rules* are in place, members of a political community²¹⁴ are legitimately entitled to abrasively blame certain offenders. We both insist that the moral license to give or take something that a person *appears* to deserve on the basis of her own individual history (be it a salary based on her work or abrasive blame based on her offense) is conditional upon the justice of a broader social practice. When the underlying practice is unjust (or in order to prevent it from becoming so), hard-workers may not be entitled to all the money they appear to deserve; it is only fair to withhold some and give it to those who lack adequate economic

²¹³ I mention Rawls’ strategy as a point of comparison to give the reader a rough idea of my strategy. The comparison to Rawls might, however, prove to be more fruitful. As critics have duly noted, Rawls maintains that people who commit crimes pre-institutionally deserve punishment whereas people who perform good work do not pre-institutionally deserve reward. Some of these critics have tried to explain away the asymmetry while others suggest that it is a deep problem for Rawls. A Rawlsian may be able to adopt the argument of this chapter as a way to address the asymmetry. I do not pursue this line of argument here, but it may be a project for future research. See Scheffler, “Justice and Desert in Liberal Theory,” and Mills, “Scheffler on Rawls, Justice, and Desert.”

²¹⁴ The members of a political community are, roughly, all the people who are citizens or residents of a nation.

opportunities. Along similar lines, as I argue in Chapter 5, members of the political community may lack the moral license to give some offenders the abrasive blame they appear to deserve for their crimes; it is only fair to spare these offenders from the political community's sting of blame.²¹⁵

Section I: Abrasive Blame in Criminal Justice Systems

Before discussing its permissibility or lack thereof, I describe the use of abrasive blame in criminal justice systems and emphasize the importance of a normative analysis of it. I focus on abrasive blame in the American criminal justice system, but I assume that the discussion can be applied to abrasive blame in other political communities' criminal justice systems.²¹⁶ The American criminal justice system, as an organ of the political community, abrasively blames the millions of criminal offenders that it incarcerates.²¹⁷ Abrasive blame is not *equivalent to* incarceration; incarceration includes deprivation of liberty that goes beyond the expression of attitudes intended to hurt a person because she did something wrong. But being abrasively blamed is almost always

²¹⁵ I have been careful to say that offenders may “appear” to deserve abrasive blame for their crimes. As I argued in Section III of the previous chapter, there is reasonable disagreement about whether persons deserve abrasive blame in the relevant sense (i.e., in the sense that would justify abrasive blame). On what I call the hard incompatibilist or the pacifist line, offenders do not deserve abrasive blame in the relevant sense. On the qualified retributivist line, offenders do deserve abrasive blame for committing crimes.

²¹⁶ In *Harsh Justice*, Whitman suggests this assumption is unfounded given the massive differences among criminal justice systems worldwide. He argues that continental criminal justice systems (especially the German system) are considerably more mild and respectful toward offenders than the American system.

²¹⁷ Sometimes offenders are abrasively blamed at the time of arrest, conviction, or parole board hearings, but prison is the most common arena for abrasive blame of offenders.

a *part* of being incarcerated in America.²¹⁸ Officials in the criminal justice system, primarily correctional officers, interact with offenders in a way that constantly expresses attitudes meant to sting them because they have committed crimes (or other infractions in the prison). The expression of these attitudes is meant to *hurt*, not physically, but certainly emotionally. This emotional hurt is hard to characterize in words, but as I did in the previous chapter, I try to describe it by offering examples and encouraging the reader to imagine what she would feel like if she were constantly subject to the expression of the same venomous attitudes.

Consider accounts from two prisoners. The first describes the remarks guards and passersby made toward Wilbert Rideau while he sat in a prison cell awaiting trial for murder: “a parade of white men, some well dressed, some not, some cops, entered the hallway periodically to stand in front of the cell and stare at me...cursing and telling me how many different ways they wanted to kill me.”²¹⁹ Jack Henry Abbott, author of the quote that opens this chapter, describes several experiences of the sting of blame in prison after being convicted of burglary. For example, consider his description of walking to the prison shower: “the guards form a loose gauntlet from your cell to the shower stall...They look at you as if you are not there...They register your facial expression to see if you are anything but meek, *humble*. Anything else raises their hackles, and their mouths turn down at the corners and they ball up their hands into fists

²¹⁸ I do not make (or reject) the strong claim that abrasive blame is a necessary part of punishment. I only claim that it happens to be a part of the American criminal justice system. As a matter of sociological fact, abrasive blame is a well-established part of American criminal justice that I do not expect to disappear anytime soon.

²¹⁹ *In the Place of Justice*, 26.

at their sides. You are nude.”²²⁰ These accounts are not descriptions of a few officials engaging in rogue behavior. The accounts describe common practice, “the norm,” in American prisons where orange day-glo jumpsuits and, increasingly zebra striped scrubs, are standard issue,²²¹ where strip searches and body cavity searches are routine, cells more closely resemble cages than rooms, and where arrangements for meals,²²² jobs,²²³ and family visits²²⁴ are shot through with a dose of derision, contempt, and disdain.²²⁵

Joe Arpaio, dubbed “the toughest sheriff in America” puts his prisoners in stripes, makes male prisoners wear pink underpants, has men, women, and juveniles work in chain gangs, and has broadcast live video footage of prisoners over the Internet.²²⁶ He keeps surplus prisoners in “Tent City” which reaches 110 degree heat under Arizona sun;

²²⁰ *In the Belly of the Beast*, 39.

²²¹ A North Carolina sheriff “noted that in 1994 he won the election by 227 votes. Then he clad his inmates in stripes. At his next election, he won by 5,000 votes, though he spent only \$15,000 to his opponent's \$100,000. ‘The public loves them,’ he said of stripes,” (Vinciguerra, “The Clothes that make the Inmate”). Some criminologists have voiced concern about the use of stripes on prison uniforms. As Charles Friel commented, “Should uniforms be the wrong size or an unusual color for the sake of putting your foot on somebody’s neck? No, I think that’s being a bully,” (ibid.).

²²² At some prisons, inmates are given less than twelve minutes to eat and some sheriffs pride themselves on moldy fifteen-cent meals (Hassine, *Life Without Parole*).

²²³ Many jobs are meaningless work for penny wages. After having been phased out, chain gangs are increasingly popular in American prisons. Compare this to the German system where offenders are allowed to have jobs outside the prison and receive normal benefits – like job protection, insurance, and four weeks paid vacation.

²²⁴ Unlike many European prisons, many American prisons only allow inmates to talk to visitors through a glass pane.

²²⁵ See Whitman, *Harsh Justice*.

²²⁶ “Con or Pawn?,” *Ottawa Citizen*.

he is infamous for telling prisoners who complained about conditions there, “it’s 120 degrees in Iraq and our soldiers are living in tents too...and they have not committed any crimes, so shut your damned mouths!”²²⁷ The general public applauds Arpaio; they enthusiastically elected him for five consecutive terms and have dubbed him a folk hero.²²⁸ The general public accepts, as Nietzsche put it over a century ago, that “being punished has something to do with being treated with derision.”²²⁹

It is important to mark a distinction between abrasive blame and degradation. Some activities in American prisons do something worse than express derision, they *degrade* offenders and attempt to strip them of their dignity.²³⁰ For example, making a person use a toilet under the eye of an official of the opposite sex (or over the Internet) does not just express an attitude meant to hurt him because he committed some crime, but it also degrades him. Some of the policies I described above are arguably degrading; they mark the offender as a person of lesser dignity (I have in mind body cavity searches without reasonable suspicion as well as broadcasting footage of offenders’ use of the toilet). In what follows, I do not mean to defend such degrading practices or the attitudes of degradation that they express. Other philosophers have done well to argue that

²²⁷ “Porridge That’s Hard to Stomach,” *Birmingham Evening Mail*.

²²⁸ Some Americans have spoken out about Joe Arpaio’s policies. Complaints and some successful lawsuits have been filed for discrimination, violations of sanitary codes, and physical violence toward criminals. Few complaints are voiced, however, over the generally abrasive atmosphere in his prison. One lawyer did describe it, however, as a “culture of cruelty.”

²²⁹ As quoted in Whitman, *Harsh Justice*, 31.

²³⁰ Whitman argues that the purpose of most practices of punishment in America is to degrade offenders – to stamp them as people with lower status.

punishments that degrade criminal offenders are unjust and impermissible.²³¹ I certainly do not stand in any disagreement with them on this point. If abrasive blame toward an offender is degrading, then it is impermissible. But I assume that a great deal of non-degrading abrasive blame goes on in the American criminal justice system. Officials in the system (and the public) express attitudes through their words and deeds that are not (at least not definitively) characterized as degrading, but are nevertheless meant to hurt offenders because they did something wrong. The serious psychological blows that this abrasive blame delivers may be equally or more difficult to justify than degrading gestures toward offenders. I want to call attention to this non-degrading abrasive blame (henceforth “abrasive blame”) and investigate whether it is morally permissible or not.

An objector may deny that officials in the American criminal justice system engage in abrasive blame toward offenders (or if they do, it is rogue behavior that is not formally or informally licensed by their office). The objector argues that what *appears* to be abrasive blame is either degrading (and so it is impermissible on grounds given above) or it is a closely related but different phenomenon that serves some justified end of the criminal justice system. Depending on what theory (or theories) of punishment the objector maintains, she might make her argument in any (or all) of the following three ways. (The response to each is roughly the same.)

First, she might argue that officials ought to act in certain ways in order to send offenders a message that their crimes were wrong.²³² Acting in these ways may look like

²³¹ See Murphy and Hampton, *Forgiveness and Mercy*; Hill, “Treating Criminals as Ends in Themselves;” von Hirsch, *Principled Sentencing*.

²³² See Feinberg, “The Expressive Function of Punishment.”

abrasive blame, but it is not. To abrasively blame is to express attitudes that are meant to sting above and beyond what is needed to make a person recognize that she did something wrong. Officials do not intentionally go above and beyond in this way.

Second, she might argue that officials ought to act in certain ways to vindicate the political community's scheme of rights; to stand up for the victim's rights which the offender violated.²³³ Abrasive blame, however, is the expression of attitudes that are meant to hurt above and beyond the amount needed to vindicate rights. Again, officials do not intentionally go above and beyond in this way.

Third, the objector might argue that officials of the criminal justice system do not intentionally express any attitudes whatsoever through the practices I have described. Officials are simply trying to maintain a reasonable degree of security within prisons where offenders must be kept to provide a reasonable degree of security within society at large.²³⁴ Doing so requires acting in ways that may appear to express attitudes of abrasive blame, but once again, it is a mere appearance (or perhaps an unintended side effect). For example, wardens issue striped uniforms to make sure prisoners can be caught if they escape, not to sting them with blame.

The objector is trying to shine benign light on abrasive blame in the criminal justice system, to explain it away as part and parcel of justifiable ends of punishment. Even if one agrees²³⁵ with the objector that officials of the criminal justice system should

²³³ See Hampton, *Forgiveness and Mercy*; Morris, "Persons and Punishment," Ripstein, *Equality, Responsibility, and the Law*.

²³⁴ Deterrence theorists may be inclined to put the objection in these terms.

²³⁵ The reader is free to press the objector on these points, but doing so is beyond my present scope.

make offenders understand they did something wrong, vindicate rights, and prevent offenders from threatening security, the objection is nevertheless unsuccessful. The objector is missing (or is blind to) facts about existing systems of criminal justice, at least the American system. Officials in the American criminal justice system express much more derision than is required to accomplish what the objector takes to be justifiable ends of criminal justice.²³⁶ The criminal justice system is replete with the expression of venomous attitudes that are superfluous to the achievement of these ends.²³⁷ For example, even if stripes are necessary for security (and I highly doubt this is the case), pink underpants underneath the stripes are certainly not necessary for security.

The presence and extent of the sting of blame in the criminal justice system may not be readily apparent to most Americans; abrasive blame, especially with respect to criminals, is “the norm” and we are sometimes prone to assume that what is normal is

²³⁶ They are certainly not alone. As Bentham observed centuries ago in England, “Legislators and men in general are naturally inclined [to err on the side of giving too much punishment]: antipathy, or a want of compassion for individuals who are represented as dangerous and vile, pushes them onward to an undue severity,” (“The Principles of Penal Law,” Part I, Ch. 6, p. 401).

²³⁷ It is worth contrasting my position more carefully to those of philosophers who argue that *condemnation* is a necessary feature of punishment (see Feinberg, Duff, von Hirsch, Hampton, Morris, and Kadish. Also see Sher, but he argues that “blame” rather than condemnation is necessary for punishment.) For the most part, these philosophers define condemnation such that it is not a species of what I am calling abrasive blame (for example, Hampton generally characterizes condemnation as the expression of attitudes that are necessary to assert that the offender’s act was wrong and/or to vindicate the system of legal rights). Put another way, I am defining abrasive blame in such a way that it is not what these philosophers call condemnation. I am not necessarily in disagreement with any of these philosophers. I am only disagreeing with someone who claims that *all* the non-degrading, but nevertheless hurtful attitudes directed at offenders in criminal justice systems count as what the above philosophers call condemnation. I think some of these attitudes are not what these philosophers call condemnation, but rather what I call abrasive blame.

required. A comparison to a different criminal justice system is helpful to shake us out of this assumption. The German criminal justice system meets justifiable ends of criminal justice as well as the American system, but it does so without very much abrasive blame. The conditions of incarceration in Germany are required by law to meet “the principle of normalcy:” life in prison should mimic life in the outside world as closely as possible. Contrast this approach to Joe Arpaio’s unabashed aim to “run a very bad jail.”²³⁸ Whereas American correctional officers call prisoners by number or last name, German officers (who are trained extensively as civil servants) address inmates respectfully as “Sie” (the polite and formal term for “you”) or “Herr So-and-So.” Moreover, laws protect inmates from being insulted by both guards and one another. Invectives cannot be hurled back and forth through bars as they so often are in American prisons. Bars on doors have been eliminated in German prisons and cell doors remain unlocked during the day. Prisoners have meaningful jobs supervised by employers rather than guards (and are even given four weeks of paid vacation!). In France, prison uniforms have been eliminated. German prisoners earn the privilege of wearing normal clothes if they have shown that they do not pose an escape risk.²³⁹

For a few years, German law required prison officials to knock upon entering a prisoner’s cell (the law no longer requires a knock, it is left to the guard’s discretion). The rationale was that “managers were to work to create suitable forms of polite interaction between guards and inmates...Entering a place of confinement without first

²³⁸ Whitman, *Harsh Justice*, 62.

²³⁹ See Whitman, *Harsh Justice* for a full description of German prison conditions compared to American prison conditions.

knocking means imposing limits on prisoners which are not necessary as an indispensable consequence of the deprivation of liberty.”²⁴⁰ The German system works to keep prison life as similar as possible to normal life; all aberrations must be justified on grounds that they are indispensable in fulfilling justified aims of punishment. Note the stark contrast to the American criminal justice system. The presumption here is to keep prison life harsh; any practice that counts as part of “prison management” is acceptable as long as it does not violate the court’s interpretation of the Eighth Amendment against cruel and unusual punishment.²⁴¹ This means that punishment cannot offend the American population’s “evolving standards of decency.” Whipping and other methods of *physical* violence toward inmates currently offend the American population’s standards of decency, but the emotional sting of abrasive blame does not. Derision does not peel off skin or dignity, so it is fair game – at least in America and in its criminal justice system.²⁴²

Abrasive blame by officials in the American criminal justice system is continuous with the harsh sentiments that the American public tends to have toward crime.²⁴³ As

²⁴⁰ Ibid., 90.

²⁴¹ In contrast to European courts, the U.S. Supreme Court “has taken a distinctly hands-off attitude toward overcrowding, observing simply that ‘restrictive and even harsh’ conditions ‘are part of the penalty that criminal offenders pay for their offenses against society,’ (Whitman, *Harsh Justice*, 61).

²⁴² As Judge Reinhardt put it in a case challenging the practice of hanging, Americans seem to have forgotten that “cruelty does not necessarily involve [physical] pain,” (Campbell v. Wood, 18 F.3d 662 (1994)).

²⁴³ I call these “harsh sentiments” instead of abrasive blame because in many cases, people who give words to these sentiments do not expect their remarks to fall upon the ears of criminal offenders and sting them (although perhaps I am being too charitable to the public here).

sociologist David Garland calls to our attention, the “openly avowed expression of vengeful sentiment” toward crime and criminals is commonplace in the American public square. In France and other more bureaucratized cultures, such avowed expressions are unacceptable. Calling a law “three strikes and you’re out,” argues Whitman, would be unheard of in French or German cultures where such “explicit attempts to express public anger and resentment are...taboo...rather than a recurring theme of the rhetoric that accompanies penal legislation decision-making.”²⁴⁴

The American court of public opinion does not save its abrasive blame for criminals. Editorial sections of newspapers are frequently ripe with invectives that are intended to sting dishonest leaders, greedy bankers,²⁴⁵ and lazy leeches. Or consider Bob Dylan’s abrasive blame toward warmongers: “for threatening my baby, unborn and unnamed, you ain’t worth the blood that runs in your veins.”²⁴⁶ Despite the pervasiveness of this kind of abrasive blame and the normative questions it raises, it is not my primary focus here. I limit my discussion to the especially powerful and consequential form abrasive blame takes when political communities collectively and

²⁴⁴ Garland, *Culture of Control*, 13-14, 41, as quoted in Whitman, *Harsh Justice*, 199-200.

²⁴⁵ One angry editorialist seems to be calling for more systematic abrasive blame toward Wall Street:

What will it take before someone on Wall Street is symbolically punished? Not just Bernie Madoff, who actually did something illegal. I mean someone who operated within the letter of the law, but outside its spirit – former treasury secretary Henry Paulson, say, for liberating banks from the need to back their loans with assets, or Merrill Lynch’s John Thain, for redecorating his office while his employees lost their jobs, or...fill in the name of the Wall Street yegg you love to hate. (Brown, “It took a jester to point the finger,” *The Globe and Mail*)

²⁴⁶ Dylan, “Masters of War.”

institutionally express it in their systems of criminal justice. It is important to recognize, however, that abrasive blame among members of the political community is not a phenomenon isolated within and tailored for use as part of criminal punishment. Abrasive blame is a widespread social phenomenon among members of a political community and it rears its head in an especially focused and powerful way in criminal justice systems. Now the question is: is such abrasive blame morally permissible? Is the American criminal justice system, in contrast to most of the German criminal justice system, *wrong* to engage in abrasive blame toward criminal offenders; or is it just different, dare I say “tougher”?

Section II: The Permissibility of the Political Sting

To investigate the permissibility of abrasive blame in a political community (and particularly in its criminal justice system), let us recall the argument presented in the previous chapter. If an act cannot be justified on grounds that one can expect people to accept insofar as they are reasonable, then all other things equal, the act is morally impermissible. I argued that absent special social conditions, the respective weight of reasons for and against abrasive blame is very well balanced. Given this fact, abrasive blame cannot be justified on grounds that one can expect people to accept simply insofar as they are reasonable. In the context of a personal relationship, however, things might be different. When parties of a personal relationship have jointly accepted an abrasive blame norm that allows abrasive blame, they can expect each other to accept a justification of abrasive blame insofar as they are reasonable people *and* reasonable parties of the relationship. Under such conditions, it is permissible for the parties to abrasively blame one another in accord with their jointly accepted norm.

Modern political communities are generally different from personal relationships in a respect that renders the above line of argument inapplicable. Although modern political communities have developed norms about when, how, and to what degree abrasive blame is appropriate, it is doubtful (at best) that all members jointly accept these norms. Most modern political communities are too heterogeneous and diffuse for us to expect that all individual members share and affirm these norms as guides to the way that they should treat one another. Individual members of the community may all be able to report what the norms are, but it is unlikely that they will all be mutually committed to the community's rules as picking out the right way of guiding behavior. For example, there are some qualified pacifists in the American political community who do not accept American social rules that deem abrasive blame appropriate. I suspect there are also some members of the German political community who object to the seemingly strict restrictions that German social rules place on abrasive blame in response to crime. Borrowing a phrase from Rawls, the "fact of reasonable pluralism" reigns with respect to abrasive blame in political communities. If abrasive blame is to be justified in a modern political community, the justification must proceed without relying on a fact of shared acceptance of rules in the community.

In what follows, I argue that abrasive blame can be justified in a modern political community in a different fashion. First, I maintain that political communities have developed *social rules* about when, how, and to what extent abrasive blame is appropriate. Then I argue that individual members have presumptive reason to follow those rules whether they accept them or not. When the rules deem abrasive blame appropriate, the presumptive reason to follow the rules *generally* tips the balance in favor

of abrasive blame. In these cases, members of the community can expect each other to accept a justification of abrasive blame insofar as they are reasonable members of the community.²⁴⁷ In the context of the political community then, abrasive blame in accord with the community's social rules is often – but, as I explain in the next chapter, not always – permissible.

The Rules

My argument begins by pointing to the fact that most political communities have community-wide social rules about when, how, and to what degree abrasive blame is allowed among the members. First, I need to briefly explain what *community-wide social rules* are.²⁴⁸ The rules are practiced standards for correct and incorrect behavior (in my case expressions of abrasive blame) within the political community. Although they are not necessarily written down or officially promulgated, members of the community generally know what they are. In order to be social rules, it must be the case that members of the relevant community generally follow the rules. They need not be entirely compliant with the rules, but they must be generally compliant and usually criticize those who deviate from the rules. People need not (although they might) accept the rules in the sense of having undertaken a commitment to follow them; they may begrudgingly follow them to avoid criticism from others. For this reason, I am using the term “social rule”

²⁴⁷ The commission of a wrong action is the “first-order” reason why a person engages in abrasive blame. That is, the first answer to “Why are you abrasively blaming her?” is “because she did something wrong.” This first-order answer is different from a deeper moral justification of abrasive blame. The deeper moral justification answers the question, “Why does the fact that she did something wrong get to *count* (morally) as a reason to abrasively blame her?”

²⁴⁸ My explanation draws on Postema, “Custom in International Law,” and his account of practiced norms of a community.

here rather than “norm” (which I used in discussion of personal relationships). “Norm” tends to connote that those over whom it applies have accepted or committed to it, whereas “rule” (or social more) tends not to have this connotation.

In order for people to have a community-wide social rules, they must not just follow the rules as a matter of what they take to be independent, personal policy. Instead, they must recognize that the same rules govern the other members of the political community. They must recognize, for better or worse, that following the rules is “how *we* do things around here.” Examples of community-wide social rules include the basic standards of etiquette, some standards of fashion, the rules of the road and even the rules of games as simple as tic-tac-toe. Smaller communities like athletic teams or orchestras have social rules governing how practice or rehearsal is to be run.

Many communities (ranging from sports associations to political communities) have community-wide social rules about when, how, and to what degree abrasive blame is appropriate. I henceforth refer to such a set of rules as a community’s abrasive blame rules or (CAB). The contours of CABs differ from community to community (consider the difference between a neighborhood full of New Jersey housewives and an academic department in Oxford), but most²⁴⁹ CABs allow for some controlled expression of abrasive blame in response to wrong action.²⁵⁰

²⁴⁹ In most communities, rules that prohibited any and all abrasive blame would be unstable, unpracticed, and would devolve into a condition without any abrasive blame rules. However, such rules may work in communities where members are socialized in a way that minimizes the disposition to abrasively blame. Quaker communities, Gandhi’s Ashram, and monasteries are arguably examples of groups that have rules against all abrasive blame. These groups seem to be exceptional in their categorical stance against abrasive blame (and indeed they all involve socialization that is distinct from mainstream

Most political communities have a CAB that consists of a large number of rules about when, how, and to what degree abrasive blame is appropriate. Many of these rules deem it appropriate for individual members to abrasively blame one another for various wrongs. For example, the American political community's CAB deems Bob Dylan's abrasive blame toward warmongers correct. It does not, however, deem it appropriate for a stranger in the community to abrasively blame me for telling a lie to friend. My focus here, however, is on a very particular subset of rules in the political community's CAB.

The CAB in most political communities has rules according to which it is appropriate (sometimes even required) for the criminal justice system, as an organ of the whole political community, to abrasively blame criminal offenders. The description of the American criminal justice system in the previous section supports this claim. Most officials in the American criminal justice system do not abrasively blame offenders willy-nilly; they abrasively blame offenders according to social rules which call for certain forms of abrasive blame in certain types of situations, but preclude abrasive blame in other forms and types of situations. These rules are generally informal norms, but sometimes they take the shape of formal policy in penal institutions and they can even be ensconced in law. According to these practiced standards, it is correct for correction officers to lock a criminal offender behind bars with disdain, to put him in stripes and strip search him regularly, to command where and when he can walk, sit, or spit. It is

culture). *Most* communities have shared practiced rules that allow for controlled abrasive blame when certain conditions are met.

²⁵⁰ CABs also seem to allow for controlled expression of abrasive blame in response to actions that reveal bad character traits. My focus, however, is on abrasive blame in response to wrong action.

incorrect for private citizens to do these things and it is incorrect for wardens to express derisive attitudes toward offenders using racist or sexist comments, physical violence, or starvation diets. Members of the American political community, officials and non-officials, recognize that these are the rules guiding how criminal offenders are to be abrasively blamed around here.

I assume that most other modern political communities similarly have a CAB that includes rules regarding when, how, and to what degree abrasive blame can be expressed toward criminal offenders. It is important to note, however, that the CABs in other communities may allow less (or more) abrasive blame than the American CAB in different types of situations. Just as the norms that parties of personal relationships jointly accept differ greatly from one relationship to another, so too do CABs differ from one political community to another.

Why Follow the Rules?

According to the CAB of most political communities, the commission of certain wrongs makes it appropriate for members of the community to abrasively blame the wrongdoer. And the commission of a crime makes it appropriate for a political community's criminal justice system to abrasively blame the offender. But the fact that there are rules according to which abrasive blame is appropriate does not entail that one has a morally legitimate reason to follow these rules. Communities in the American South had rules according to which slavery was appropriate, but that did not entail that Southerners had a morally legitimate reason to follow those rules. If I am to show that abrasive blame is morally permissible in political communities (and particularly in criminal justice systems), I need to show not just that political communities have rules

about abrasive blame, but that people have moral reason to follow those rules. What gives the CAB its reason-giving or normative force (insofar as it has any) in the political community?

In a full theory about the justifiability of abrasive blame, I would also discuss the justification (or lack thereof) of abrasive blame in communities other than the political community. Sometimes abrasive blame can be justified in communities through the kind of argument I gave in the previous chapter (perhaps the community is a many-membered personal relationship in which all parties share acceptance of an abrasive blame norm). In other contexts, members of the community may not have a *shared acceptance* of a blame norm, but they might have a CAB that has reason-giving force. For example, this might be the case in an academic department. What the reason-giving force of the CAB is (if it has any at all) will highly depend on the type of the community. In what follows, I focus on why the CAB of *the political community* has reason-giving force.²⁵¹

²⁵¹ One might object that in order to investigate the permissibility of abrasive blame in the criminal justice system, we should look at the CABs in the smaller “communities” of individual prisons in America (and elsewhere). I reject this move because the criminal justice system, including its prisons, is an organ for the *political community*’s expression of abrasive blame. The derision that offenders receive from correctional officers within prisons is allowed, often encouraged (as discussed above, Sheriff Arpaio is a “folk hero” who is consistently re-elected by the public), and authorized by the political community. The norms that govern this expression are part of the political community’s CAB about when, where, who, and how people can engage in abrasive blame. If wardens violate these norms, they might be criticized by other guards or prisoners, but what is important for my purpose is that they are criticized by the political community as violating *our* standards of decency about blame (consider, for example, the public reaction to Abu Ghraib). As Whitman illustrates, harsh treatment (including abrasive blame) in American prisons is continuous with, not an aberration from, American culture (including its abrasive blame norms). It is within the political community’s power to stop or significantly curb back expressions of abrasive blame on a wide scale (as Germany has done), but this is very unlikely to happen without change in the political community’s CAB. Moreover, the enterprise of justifying abrasive blame from within specific prison

In what follows, I argue that members of the political community have good *presumptive reason* to follow their CAB as long as their CAB is not unjust. The argument has four basic steps. First, there is good reason for a political community to have social rules about how, when, and to what degree abrasive blame is appropriately expressed. People naturally engage in abrasive blame; that is, they are psychologically predisposed toward abrasively blaming those whom they judge to have done wrong. Moreover, in human groups (especially large ones) abrasive blame tends to cycle upwards in severity if left unchecked – derision tends to breed more intense derision which tends to breed hostility. If political communities had no CAB, then people’s natural disposition to abrasively blame would lead to dangerous instability. Political communities would be vulnerable to a spiral of costly vendettas; too much time would be spent blaming rather than engaging in other, more fruitful, pursuits. Without any CAB, people’s chance of living together peacefully in their political communities would significantly decrease. It would be more likely that life would be “nasty, brutish, and short” in human political communities. CABs help prevent this volatile state of affairs by allowing for controlled expressions of abrasive blame that are less likely to snowball; CABs provide an important social check on people’s blaming behavior.²⁵² Whether CABs historically developed as a response to this need is an anthropological question

“communities” is likely futile. Given that the “community” consists of guards and prisoners kept in cells, I suspect that any CAB that is specific to this group but not to the political community at large would have very little reason-giving force (certainly not enough reason-giving force to justify the guards in abrasively blaming the prisoners). Since I do not, however, discuss how abrasive blame is justified in smaller communities, I can do nothing more than register my suspicion here.

²⁵² My argument here loosely echoes that of John Stuart Mill and Adam Smith. Both argued for the need to civilize resentment with some kind of social rules.

beyond the scope of this paper. My claim here is simply that the presence of some CAB in a political community – regardless of how it arose – significantly increases the chance that life in the community will be stable and peaceful. I assume that we have good reason to seek and maintain peace in our political communities, so we have good reason to have some CAB that increases our chance at realizing such peace.

This claim stands in need of qualification. The pursuit of stability does not count as a reason in favor of having *any* social rules about abrasive blame in the political community. It does not count in favor of having unjust rules about abrasive blame. Remaining consistent with Part I of the dissertation, I call a rule unjust if it is incompatible with a full theory of political justice. Although the exact relationship between justice and the pursuit of stability is contested,²⁵³ it is clear enough that we do not have morally legitimate reason to pursue stability at the price of supporting injustice. (A grossly unjust fascist dictator might increase a community's chance for stability but that is not a morally legitimate reason to have such a dictator!) A political community has moral reason to have rules about abrasive blame, but only insofar as those rules are not unjust (that is, insofar as they are decent).

One might wonder what standard of justice is at work here. As I argued in Chapter 3, there is reasonable disagreement from the point of view of comprehensive moral theories on whether abrasive blame is morally permissible. A social rule is unjust not because it is ruled out by a comprehensive moral theory, but because it is ruled out by

²⁵³ Is the maintenance of stability a proper aim of justice (as Rawls suggests at times) or is it an aim that is at odds with justice? Rousseau, unlike Rawls, holds that it is impossible for a fully just institution to remain stable. Investigating this debate is beyond my present scope.

a theory of political justice. The shared political value of treating persons as free and equal persons puts constraints of political justice on rules of abrasive blame. For example, the rules about abrasive blame should not be racially biased and they should respect individuals as free and equal persons. In most political communities, it is likely that there will also be constraints on abrasive blame that arise from an overlapping consensus of comprehensive moral views in the political community. Despite significant disagreement about abrasive blame among the four positions I described in Chapter 3, there is still some shared ground that can be appealed to in arguments from public reason. For one thing, all of these views stand against extremely harsh expressions of abrasive blame in most cases. Both of the reasonable positions in favor of abrasive blame, the qualified retributivist and the tough-love pacifist, recognize upper bounds or at least some constraints on the amount of abrasive blame that is permissible (the qualified retributivist recognizes a proportionality constraint and the tough-love pacifist recognizes an efficiency constraint).

A CAB is a set of many practiced rules about abrasive blame and it may include rules that are unjust and others that are decent (i.e. not unjust). In the next three steps of the argument, I show that members of the political community have presumptive reason to follow the decent rules in their CAB (and lack a morally legitimate reason to follow the unjust rules of their CAB). To simplify the writing of the argument, I will speak generally rather than strictly. I use the term DCAB (“set of decent community abrasive blame rules”) to refer to a CAB that consists of rules that are, for the most part, not unjust. And I will argue that members of a political community have presumptive reason to follow a DCAB.

With this first step and its qualification behind us, let us turn to the second step of the argument. Due to the fact that there is reasonable disagreement on when abrasive blame is reasonable, there are a wide variety of DCABs that a political community might have. Although there is a diversity of options, a political community actually practices one DCAB, its own. A political community's own DCAB is the set of norms that its members actually share and practice about how, when, and to what degree abrasive blame is appropriately expressed in the community. Members may reasonably maintain that their political community's DCAB is suboptimal – it is not the one that they would have in place given their own personal and reasonable convictions about abrasive blame. Insofar as they are reasonable, however, they must all recognize that their political community's DCAB has the distinct advantage of existing as a practiced rule of the community. Regardless of whether they agree that a community's actual DCAB is ideal, they must recognize that it fills a morally important niche. As explained above, a DCAB helps the community seek the worthwhile aim of peace and stability.

Third, the fact that a DCAB fills a morally important niche in a community gives members a good reason to sustain the DCAB. Recall that the social rules that make up the DCAB require general compliance in order to persist. If members were not to generally comply with these social rules, then the rules would cease to persist and the community members would have failed to sustain their DCAB (a result they have reason to avoid). So members have reason to generally comply with their DCAB. More specifically, they have presumptive reason to follow their DCAB: a reason that generally and non-decisively counts in favor of following their DCAB. The reason does not count in favor of following the DCAB in every instance; it only counts in favor of following the

DCAB frequently enough to sustain *general* compliance. Also, the reason to generally comply with the DCAB is non-decisive; it can be outweighed by other strong considerations in particular cases.

The pieces are now in place to tie the argument together. Most modern political communities have decent social rules that deem it appropriate for members to abrasively blame one another for certain wrongs and for the criminal justice system to abrasively blame offenders for crimes. The rules (the DCAB) are worth maintaining because they are a reasonable source of stability for the community. As such, there is a presumptive reason for members to follow these rules.²⁵⁴ Barring special countervailing reasons in particular cases, members of the political community can reasonably expect one another to accept a justification for abrasive blame that is in accord with the DCAB.

Accordingly, abrasive blame in accord with the DCAB is *prima facie* morally permissible.

Members of most modern political communities are not like members of a personal relationship who together carve their own paths through morally indeterminate terrain, but this does not mean they have no path through this terrain. They are like members of a herd that has developed a relatively clear migration course over time.

Individual members often have reason to stay the course – to follow the rules that have

²⁵⁴ We should be careful to note that although the presumptive reason to follow the DCAB is crucial to the justification of abrasive blame in the political community, it is usually not the reason *why* members abrasively blame one another. Members blame a person because she has done something wrong (or committed a crime). My argument here explains why wrongdoing (or crime) usually counts as a morally legitimate reason for expressing abrasive blame within the political community.

been developed. But sometimes, as I discuss in the next chapter, they should turn away, hoping (but not necessarily expecting) the herd to follow and strike a better course.

CHAPTER 5: SUBSTANTIAL WOUNDS AND THE BANDAGE OF MERCY

“Until you put these things right, you’re not entitled to boast of the justice meted out to thieves... You allow people to be brought up in the worst possible way, and systematically corrupted from their earliest years. Finally, when they grow up and commit the crimes that they were obviously destined to commit, ever since they were children, you start punishing them. In other words, you create thieves, and then punish them for stealing!”

~Sir Thomas More²⁵⁵

In the previous chapter, I showed that some political communities abrasively blame criminal offenders through the organ of the criminal justice system. I explained how such abrasive blame might be *prima facie* morally permissible. Insofar as *decent* community-wide social rules deem it appropriate for the criminal justice system to abrasively blame offenders, it has presumptive reason to do so. If this presumptive reason is *not undermined or outweighed*, it is permissible for the criminal justice system to abrasively blame criminal offenders in accord with the rules. Although my own convictions about abrasive blame make me refrain from applauding “America’s toughest sheriff” Joe Arpaio, I have argued that his abrasive blame toward criminal offenders might be permissible. It depends, of course, on whether his policies are in accord with American social rules (which they seem to be given widespread popular support), whether they are decent (at least some seem to be compatible with a theory of political

²⁵⁵ *Utopia*, 27.

justice), and whether the presumptive reason to follow the rules is washed out by any stronger considerations in play.

My task in this chapter is not to articulate all the social rules governing abrasive blame of criminal offenders and determine whether they are reasonably defensible. Nor is it my task to determine whether the presumptive reason to follow the rules is outweighed or undermined by other considerations in individual cases. I cannot determine the all-things-considered permissibility of abrasive blame in criminal justice. To do this, I would need to do substantial policy research on existing systems of criminal justice. Instead, I am highlighting a *general deficiency* in social rules that govern abrasive blame toward criminal offenders in several political communities. In several of these communities (e.g. America), the rules deem it appropriate for the criminal justice system to abrasively blame what I call battered and fragile offenders just like other offenders. In Section I and II, I argue that it is unfair for the political community to abrasively blame battered and fragile offenders. I argue in Section III that criminal justice officials have strong reason not to abrasively blame these offenders and to change social rules that would license such blame. In general, abrasively blaming battered and fragile offenders usually does not sting them permissibly; it unfairly strikes at them and inflicts substantial wounds.

Section I: Battered Offenders

In the next two sections, I describe two types of offenders: battered offenders and fragile offenders. I assume that some political communities have social rules according to which it is appropriate for the criminal justice system to abrasively blame these

offenders for committing crimes. I argue that it is unfair, however, for the political community to abrasively blame these offenders.

I begin by defining a battered offender as a person who 1) has been denied rights, 2) has culpably violated criminal law, and 3) would likely not have committed the criminal offense had her rights been satisfied. She is a person for whom the denial of rights played a causally significant role in her culpable commission of crime.

Identifying persons as battered offenders crucially depends upon identifying the rights that people have. I will assume that people have the following rights (the list is not exhaustive): rights to basic liberties,²⁵⁶ to fair equality of opportunity, to physical security, and to economic security required for a decent quality of life.²⁵⁷ In most political communities, there are at least some people who are denied these rights. Of these people, most abide by criminal laws or at least do not culpably violate them. Some culpably violate criminal laws in ways that are not significantly connected to their denial of rights. Others, however, culpably violate criminal laws in ways that are causally linked to their denial of rights. I assume that these people exist and I call them battered offenders.

As I define the battered offender, she is culpable for violating criminal law; she is fully responsible for violating the law and none of the usual excusing or justifying conditions apply to her. The battered offender is importantly distinct from a person whose rights are denied in a way that leads her to excusably or justifiably act in a way

²⁵⁶ More specifically, people have a right to “a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all,” (Rawls, *Justice as Fairness*, 42).

²⁵⁷ It is beyond the scope of my project to defend this list of rights.

that is not in accord with criminal law.²⁵⁸ For example, consider a person who is acutely denied the right to physical security. He is severely abused in such a way that he loses the ability to control himself from assaulting those who remind him of his abuser. This person meets a man who reminds him of his abuser and he is unable to restrain himself from assaulting the man. This person is not a battered offender; he should be excused on grounds of diminished responsibility. To take another example, consider a person who is acutely denied her right to economic security. She is on the brink of dying of starvation and decides to steal food (or the necessary means to acquire it). This person is not a battered offender; she is a person whose act was justified on grounds of extreme necessity.²⁵⁹

Unlike these individuals, a battered offender is fully culpable for a violation of criminal law. In practice it is very difficult to determine whether an offender is a battered offender. One must not only confirm that she has been denied rights, but also confirm a) that the denial has been causally significant to her commission of the crime and b) that she is nevertheless responsible and for an unjustifiable offense. Although the task of identifying battered offenders is difficult, the criminal justice system can be expected to perform this task. In other arenas, it is already charged with the task of identifying what factors count as causally significant to human action as well as determining agents'

²⁵⁸ On one reasonable interpretation of the quote from Thomas More that opens this chapter, More is arguing that disadvantaged criminals are actually not culpable. I do not think, however, that this is the only available interpretation and it is not the one I prefer.

²⁵⁹ The French and German penal codes excuse homicide on these grounds (French Penal Code art. 64; German (Federal Republic) Penal Code sect. 35).

responsibility and the presence of excusing or justifying conditions. I offer the following examples of battered offenders:

Underprivileged Offender: Anna is a young woman who grew up in foster homes after her father left and her mother died from a drug overdose. Although she attended school for twelve years, she never learned how to read or to do addition past a first grade level. Her neighborhood has always been full of gangs, one of which she joined for protection. She has never known anyone who could be described as a positive role model. The only family she knows is an older brother who is in prison. At age nineteen, Anna is impoverished; from where she stands, she can see no job prospects or positive direction in which to take her life. Like many of her friends, she becomes more involved in her gang. When she is twenty, Anna is convicted of prostitution, shoplifting, and possession of drugs with intent to sell. Anna is fully culpable for her criminal offenses (she knew they were wrong, she had the self-control to refrain from committing them, she was not facing conditions of extreme necessity). But she has also clearly been denied a right to fair equality of opportunity. If she had received a fair equality of opportunity, it is very likely that she would not have engaged in criminal activity.²⁶⁰

²⁶⁰ Contrast Anna to Johnny. Johnny's childhood was roughly the same as Anna's, but he educates himself at a library downtown, goes to college, and becomes a broker on Wall Street. Although he earns an excellent salary, he commits and is convicted of massive fraud. He too was denied fair equality of opportunity. The denial was not, however, causally significant to his commission of crime. Johnny is not a battered offender.

Abused Offender:²⁶¹ David is a man who suffered severe physical and psychological abuse as a child. His father abandoned the family and his mother neglected and abused him. She would lock the refrigerator and only reward the children with food when they stole for her. She hosted a motorcycle gang in her home that sadistically tortured David and carried out sexual acts in front of him. The state removed David from his mother's custody when he was fourteen and rotated him through a series of inadequate foster homes where he was subject to more abuse and neglect. When David was nineteen, he stabbed a man to death during a botched burglary. The criminal justice system deemed David fully responsible and culpable for his offense; he knew right from wrong and had enough self-control to refrain from committing murder. I am not sure whether the criminal justice system was correct in judging that David was fully culpable for his offense, but let us assume that he was for the sake of the argument here. His right to basic security was denied for the fourteen years in which he incurred severe abuse. It is very likely that he would not have committed murder if he had not been abused.²⁶²

²⁶¹ Unlike the previous example, this one is not hypothetical. It is based on David Woods who was executed in 2007 for the murder of Juan Placencia (Woods v. McBride, 430 F.3d 813, 824 (2005)). Other examples of offenders like David Woods include Robert Harris (although he may not have been fully culpable; see Watson, "Responsibility and the Limits of Evil") and Wilbert Rideau, *In Place of Justice*.

²⁶² Contrast David to Gladys. Gladys is married to a man who beats her. She lacks the resources to leave him (she is not empowered to do so and she fears her husband will track her down and the police will not protect her). After several years together, her husband leaves her. Before, during, and after the years she is with her husband, Gladys participates in smuggling illegal immigrants. She is convicted for her offenses. Although her basic right to security was denied during her time with her husband, the denial was not causally significant to her crime. Gladys is not a battered offender.

Unjustly Punished Offender: Henry is a member of a racial minority and commits a minor theft. He has no other criminal record and no history of violence, but a racially prejudiced judge gives him an unjustly long sentence of twenty-five years. He is angry about having to serve such a long sentence and his anger often lands him in solitary confinement. He is beaten by guards, raped by other prisoners, and eventually joins a gang in prison for protection. When he is released, Henry cannot find a job; no one wants to hire a fifty-year-old ex-convict with no job experience aside from menial labor in prison. He has lost any respect for the law that he once had. Making use of his connections in the gang, Henry becomes a hardened criminal. He commits and is convicted of armed robbery and assault. He was responsible for his crimes, but it is very likely that he would not have committed them if he had not been denied his rights to liberty, security, and fair equality of opportunity.

I argue that the political community has strong moral reason to refrain from abrasively blaming battered offenders even though the DCAB calls for such blame. First I sketch the arch of my argument and then I defend its core claims. The argument begins with the claim that the political community is collectively responsible for the battered offenders' denial of rights. Given that this denial plays a causally significant role in the crimes of battered offenders, the political community is in some sense *complicit* in the battered offender's violation of criminal law. Although the battered offender is fully responsible for her violation of criminal law, the political community bears some responsibility as well. The community must accept responsibility for its own faults that

played a crucial role in causing the offense. This acceptance of responsibility is incompatible with abrasively blaming a battered offender.

Defense of this line of argument hangs on defense of the claim that the political community is complicit in the battered offender's violation of criminal law. As a collectivity, the community bears responsibility for the denial of rights that was crucial in causing the battered offender's violation of criminal law. The community does not bear collective responsibility for each individual act (or omission) that violated the offender's rights. The community as a whole, however, is responsible for the fact that the offender's rights were denied – that is, systematically violated with general impunity.²⁶³ When the battered offender's rights were violated, the community did not make a concerted effort to vindicate her rights, make amends, or take steps to assure that her rights would not be violated over and over again. To systematically fail to take adequate steps upon violation of a person's rights over long periods of time (decades, but maybe less) is to deny them those rights. As Arthur Ripstein puts the point, "a right that can knowingly be violated with impunity is no right at all."²⁶⁴ This systematic failure to stand up for the rights of battered offenders is a collective wrong attributable to the political community.²⁶⁵

²⁶³ The political community often prosecutes child abuse, but such prosecution does not qualify as adequate vindication of the child's right to basic security unless the child is guaranteed a safe home. It also does not qualify as vindicating the child's right to fair equality of opportunity unless the child is guaranteed help to overcome the scar left by severe abuse.

²⁶⁴ *Equality, Responsibility and the Law*, 158. It is beyond the scope of this dissertation to give an account of rights and use it to rigorously defend this claim. I am drawing on the work of others to defend this claim.

²⁶⁵ I assume here that a collectivity can be responsible for a wrong. Discussion and defense of this assumption is the scope of my project, but I direct the reader's attention to the following work: May, *Sharing Responsibility*; May and Hoffman, *Collective*

As a whole, the community is responsible for denying the battered offender her rights. That is obviously unfortunate, but why does it matter to attribute blame in criminal punishment? When a party (in this case the political community) is responsible for some wrong that plays a significant causal role in another agent's culpable violation of law, that party is complicit in the violation. We need to take a detour into the morally mysterious territory of complicity, broadly understood as "participation in the wrongs of another."²⁶⁶ To be complicit is to somehow contribute to the commission of an offense in a way that makes one partially (or fully) at fault for it, but in a way that is distinct from culpably committing the offense oneself.²⁶⁷ It is a way of being partially (or fully) responsible for an action through the mediation of another agent or agents. But how?

Complicity, like straightforward responsibility, comes in degrees. If a party acts with full knowledge and intention to assist or encourage another agent to violate criminal law, the party is complicit in a very strong sense. She is an accomplice or co-conspirator and her complicity makes her equally (or sometimes even more) culpable as the principal agent. Each member of a conspiring group of bank robbers is fully culpable for the robbery even if only one of them performed actions that actually violated law. Under the doctrine of complicity in American law, each is liable for criminal punishment for their

Responsibility; Smiley, Moral Responsibility and the Boundaries of Community; and Norrie, Punishment, Responsibility, and Justice.

²⁶⁶ Gardner, "Complicity and Causality," 127.

²⁶⁷ According to Gardner, complicity, generally, is "participation in the wrongs of another." More specifically in certain bodies of law, it is "a mode of contribution" that demarcates a wrongdoer as an accomplice rather than a principal ("Complicity and Causality," 128).

contribution to the offense.²⁶⁸ The political community is not complicit in this strong sense for the offenses committed by battered offenders. The reason is simple: the political community does not have full knowledge and intention to assist and encourage battered offenders to commit crimes.

There is, however, a weaker sense of complicity. One is complicit in a weak sense if one recklessly, albeit not intentionally, sets the stage for another person to commit a criminal offense. Consider the following examples. A man sells guns to minors whom he has heard talking about vague plans for a shooting; the minors use the guns to commit the massacre at Columbine.²⁶⁹ A person engages in a road race on a two-lane public street; his competitor swings out to pass and collides with another car, resulting in the death of the driver of the other car.²⁷⁰ A person plays Russian roulette with another in the course of which the other shoots and kills himself.²⁷¹ In *Rebel Without a Cause*, Jim plays a game of chicken in which he and his opponent (Buzz) drive stolen cars toward a cliff. Buzz's seatbelt malfunctions and he drives off the cliff to his death.²⁷² In each case a person acts in a wrongful way that unintentionally, but with reasonable foreseeability, results in another's crime. Intuitively, the person is responsible

²⁶⁸ Kadish, "Causation and Complicity."

²⁶⁹ Weisberg discusses this example of complicity in "Reappraising Complicity."

²⁷⁰ *Jacobs v. State*, as cited in Kadish, "Causation and Complicity."

²⁷¹ *Commonwealth v. Atencio*, as cited in Kadish, "Causation and Complicity."

²⁷² Another example is in *Rachel Getting Married*. A mother left her substance-abusing teenage daughter Kym home alone to take care of her kid brother. Kym is high, takes her brother to a nearby park to play, and then drives off a bridge with him in the car. He dies, Kym survives.

not only for her own wrong action (selling guns to minors, speeding, and gambling with her life), but she also bears some responsibility for the resulting crime. Her responsibility does not take away responsibility from the principal actor; the principal actor is fully culpable for making his own choice to kill, to drive recklessly, to shoot himself, and to drive full speed toward a cliff. The complicit actor sets the stage for the principal actor to get out there himself and break a leg, not to have his arm twisted.

This kind of unintended but reasonably foreseeable participation in another person's criminal act (call it "weak complicity") is often not deemed strong enough to make an actor fully culpable for the principal agent's crime. Weak complicity does not seem, at least to most, to be adequate grounds for criminal punishment for the ensuing crime. (Although the American criminal justice system did find the actor in the car race guilty of manslaughter and the actor in the Roulette game liable for involuntary manslaughter.) The jury is out on this question and it appears to be ripe for closer philosophical investigation. Fortunately, however, I do not need to answer whether weak complicity grounds full culpability or justifies criminal punishment for ensuing criminal offenses. I only argue that weak complicity grounds *some* responsibility or liability for ensuing criminal offenses. As Sanford Kadish points out by appealing to our intuitions and judicial decisions on some of the above cases, "it is both sound in policy and conformable to our intuitions of just blaming to hold a person liable for recklessly facilitating the criminal action of another."²⁷³ The law of torts offers support for this point. As Kadish explains, "if the voluntary and intentional action of another was one of the risks in virtue of which the defendant's conduct was negligent...tort liability for

²⁷³ "Causation and Complicity," 409.

damages could be imposed on the defendant.”²⁷⁴ In other words, even if weak complicity does not ground full culpability for the ensuing crime, it at least grounds civil liability for damages caused by the ensuing crime.

The example from *Rebel Without a Cause* helps illustrate this point. Regardless of whether it would be right to criminally punish Jim for the death of Buzz, it is right for Jim to take up a special sense of responsibility for Buzz’s death. Jim’s acceptance of responsibility does not in any way deny his recognition that Buzz made his own, fully autonomous choice to take the risk that led to his death.²⁷⁵ Jim recognizes that his own behavior calls out for special justification with respect to Buzz’s death. Given that he has no justification, his behavior calls for acceptance of liability and some kind of atonement for Buzz’s death. Even if it is not right to send Jim to prison, it is certainly not right for Jim to walk away and deny all liability in Buzz’s death.²⁷⁶ The point comes to the fore when Jim argues with his parents about what to do after the incident. His parents try to convince him that it is enough for him to know that he did something wrong, but that it is

²⁷⁴ “Causation and Complicity,” 403. To be clear, however, Kadish does not label what I call “weak complicity” as “complicity.” Kadish argues that these cases fit neither in the doctrine of causation law nor complicity law. The defendant was not the cause of criminal action because, at least in criminal law, one cannot be the cause of another’s commission of crime; the other is a free agent, unmoved mover (except in cases where her agency is somehow compromised). The defendant was also not strictly complicit because, for Kadish at least, complicity requires the defendant to intend the crime. He calls these kinds of cases “problem cases” for the legal doctrine of complicity and causation.

²⁷⁵ This provides evidence for Kadish’s observation: “The grip of the conception that a voluntary human action bars assigning causal responsibility to an earlier actor, pervasive as it is in the law, is loosened by the pull of holding people liable for recklessly providing others with an occasion to do harm,” (“Causation and Complicity,” 402).

²⁷⁶ For deeper analysis on the kind of failure that Jim might have if he were not to accept responsibility, see Susan Wolf, “The Moral of Moral Luck.”

silly and “idealistic” to take public responsibility for it. They advise moving away and maintaining an evasive stance with the police. After all, “nobody thanks you for sticking your neck out.”²⁷⁷ But Jim rightly rejects his parents’ suggestion. He states simply and forcefully, “I am involved... Mom, a kid was killed tonight. I don’t see how I can get out of that by pretending that it didn’t happen.”²⁷⁸ His acceptance of responsibility is palpable and intuitively correct.

Even if weak complicity does not ground criminal culpability, it certainly does ground some morally important responsibility and liability. Let us return to the context of the political community and the battered offender. Although the political community does not intend for, assist, or encourage the crimes of battered offenders, the political community is responsible for wrongs – the denial of rights – that set the stage for these crimes.²⁷⁹ The political community knows that by denying some of its members the right to security and fair equality of opportunity, it puts these people at substantially higher risk to commit crime.²⁸⁰ Given contemporary sociological analysis, we can reasonably foresee that a significant percentage of those denied their rights will go on to commit crimes that they otherwise would never have dreamt of committing. The political community ought to recognize the role it plays in the crimes of these battered offenders.

²⁷⁷ *Rebel Without a Cause*.

²⁷⁸ *Ibid*.

²⁷⁹ The chance that a young black man in America will commit a crime that lands him in prison is 1 in 8; the chance that a player in Russian roulette will commit a crime (of suicide) is 1 in 6. The analogy is not too far off.

²⁸⁰ Unlike the examples I provided above, the wrongs are actually done to the person who commits the ensuing offense. This difference only seems to heighten the political community’s complicity.

It is weakly complicit. Like Jim, it should step up and accept its responsibility and liability. Acceptance of this liability does not entail any suggestion that the battered offender is less culpable for his offense. The political community is not taking any responsibility away from the battered offender; it is rightly sharing in it.

It is difficult to discern how the political community should act in accepting responsibility for its complicit role in battered offenders' crimes. Simply knowing that the political community played a role is not enough, but what is enough? Generally, a person who bears responsibility by being weakly complicit should openly admit her fault and the role it played, apologize for it, make amends, and put forward a sincere effort to fix the underlying problem. But a collectivity like the political community may need to bear its responsibility somewhat differently than an individual would. It is beyond the scope of my project to say exactly how the political community should act in light of its complicity. Proper acceptance of its complicity may well be compatible with punishing the battered offender, but I leave this important question for future work.²⁸¹ Here I want only to raise a red flag with respect to abrasive blame.

The political community's responsibility in the crimes of battered offenders stands in strong tension with the political community abrasively blaming battered offenders. The tension presents itself in (at least) three ways. First, expressions of

²⁸¹ Dostoyevsky gives us an eloquent invitation to do this work:

A criminal can have no judge upon the earth until that judge himself has perceived that he is every bit as much a criminal as the man who stands before him, and that for the crime of the man who stands before him, he himself may well be more guilty than anyone else. Only when he grasps this may he become a judge. (*Brothers Karamazov*, trans. McDuff, 415)

How does a judge who "grasps this" respond to offenders that stand before him?

abrasive blame tend to single out the offender, separating him off from others for the sting. Consider expressions like, “you had it coming” (as if the rest of us did not?) and “you should rot in _____” (while the rest of us carry on nicely?).²⁸² This singling out of the offender seems incompatible with *sharing in* responsibility for his crime. Second, if the political community accepts its responsibility, it should own up to its denial of the battered offenders’ rights in the past. This seems to require an apologetic stance toward the offender that is incompatible with stinging him with abrasive blame. Try sincerely saying “sorry” and spitting in someone’s face. It does not work.²⁸³ Third, accepting responsibility seems to require making (or renewing) a commitment to move forward in addressing underlying faults that led to the problem. Stinging battered offenders is often in tension with this commitment. The offender should be assured that he does have equal rights in the community even if they were denied in the past. But abrasive blame does not give this forward-looking assurance; instead, it piles hurt on top of the denial of rights that offenders have already been made to endure. Abrasive blame does not help the political community disavow its prior denial of offenders’ rights, instead it seems to affirm (or at least condone) that denial.

Although some punishment may be compatible with the political community’s acceptance of responsibility in the crimes of battered offenders, abrasive blame is

²⁸² The glaring exception here is: “I’ll see you in hell.”

²⁸³ Jack Abbott puts his finger on the contradiction:
The law has never punished anyone for hurting me...Some of us prisoners are a product of prison conditions that are today recognized as ‘unconstitutional,’ indeed, criminal [horsewhipping, corporeal punishment, starvation]. What are we supposed to do? No one has yet apologized to us. The same pigs – or their stripe – still preside over these prisons. (*In the Belly of the Beast*, 117, 24).

generally incompatible with such acceptance. A political community that abrasively blames battered offenders fails to take seriously its own responsibility that it bears for their offenses. It does not act in a way that respects or aims to restore reciprocity in its relationship with the offender.²⁸⁴ Instead, it engages in a kind of hypocrisy writ large (blaming others while failing to actively accept responsibility oneself). Such hypocrisy is not only dishonest, but also *unfair*.²⁸⁵ It does not give battered offenders the apology or the future reassurance that they are entitled to as people whose basic rights have been denied. Nor is it simply silent. Instead, it singles these people out for yet another hurt (the sting) which they would likely not have received had their rights not been denied in the first place. The political community takes for itself a benefit (the license to abrasively blame) that is not fit for its “dirty hands,” while imposing yet another burden (the sting) on the already battered offender. Doing so is not fair.

One might object that the political community is acting unfairly if it abrasively blames non-battered offenders and not battered offenders who commit exactly the same offenses. The political community’s action may appear unfair at a superficial level because it treats the offenders differently without marking any distinction in their actual offenses or culpability. Treating people fairly, however, is not fundamentally about treating all like *crimes* alike, but about treating all members of the community as *free and*

²⁸⁴ This argument may suggest that the political community should not only withhold abrasive blame, but also a significant degree of criminal punishment in other forms. I have a great deal of sympathy with this suggestion and I leave it for future work. Unfortunately, it is beyond the limitations of this dissertation.

²⁸⁵ Although the collectivity has a reason to refrain from abrasively blaming the offender, this reason may not apply to some individual members of the collectivity (like the victim). The reason does apply, however, to officials acting in their capacity as organs of the collectivity (such as judges, correctional officers, and parole board members).

*equal persons.*²⁸⁶ The ideal of “treating like cases alike” requires us to identify the relevant respects in which cases are alike. Two similar cases are not “alike” in the relevant respect if the state has battered one of the offenders but not the other. This is a difference that should make a difference. The political community must treat battered offenders differently than many of those who commit the same crimes in order to treat them all fairly as equal persons in the community. An appeal to the political community’s complicity is not only *not unfair* in distinguishing one offender from another, but it is actually *motivated by a more comprehensive sense of fairness*.

Section II: Fragile Offenders

It is unfair for the political community to abrasively blame a second type of offender that I call a fragile offender. The fragile offender is a fully culpable criminal offender whose social basis for self-respect hangs in jeopardy. To describe this type of offender and to argue that the political community should not abrasively blame them, I need to explain what a social basis for self-respect is and why it is important. A social basis for self-respect is, to draw on John Rawls’ definition, the basic social “institutions that are normally essential if citizens are to have a lively sense of their worth as persons and to be able to advance their ends with self-confidence.”²⁸⁷ Self-respect is distinct from self-confidence, self-assurance, or generic warm-fuzzy feelings about oneself. Self-

²⁸⁶ I am drawing on Rawls and others who argue that fairness (and justice for that matter) is fundamentally a matter of treating people as free and equal persons. See *Theory of Justice, Political Liberalism, and Justice as Fairness: A Restatement*.

²⁸⁷ *Justice as Fairness*, 59. For Rawls, the social-basis is a feature of “basic institutions” rather than what I have loosely called the “social fabric.” He conceives of the social-basis for self-respect as a primary good. It is an important factor that parties in the original position consider when they are drawing up the principles of justice.

respect is not pride in one's actions or traits regardless of whether they are morally good or bad. Instead, self-respect is a *morally decent sense of self-worth*; a sense that one is no less of a person than anyone else, that one is a moral agent who has equal moral standing with other such agents. Self-respect is importantly distinct from the social basis for self-respect; the latter is what is normally essential in the fabric of a society in order for people to have the former. A social basis for self-respect is the kind of underlying social soil most humans need to become and remain self-respecting persons.

Having self-respect is essential for a person to develop what Rawls calls her two moral powers: the capacity for a sense of justice (to understand, apply, and act from principles of political justice) and the capacity for a conception of the good (to have, revise, and rationally pursue a conception of the good). Having self-respect is essential to act autonomously and to be a free and equal person. If members of a political community treat one another as free and equal persons, they should (among other things) design their social institutions in such a way that provides, rather than denies, each member a social basis for self-respect. The community cannot guarantee that each member will have self-respect, but it should design its institutions (including its system of criminal justice) to provide a social basis for it.

The basic institutions of many existing political communities provide a social basis of self-respect for most people. Many institutions that are supposed to provide a firm soil for self-respect instead provide a soil that has deep fissures and cracks. Many people fall through these cracks; they lack a social basis for self-respect. This leaves us in the realm of non-ideal theory. The political community should not just give up on these people who find themselves without a social basis for self-respect. Instead, it

should have policies to provide a safety net for these people. Even if it cannot ensure that these people develop self-respect, it can provide social conditions that act as a back-up social basis of self-respect. The political community should do what it can (within reason) to provide all members with an adequate social basis for self-respect – be it through its basic institutions or supplementary safety nets.²⁸⁸ To fail to do so is to fail to make reasonably available the tools that human beings need to have a sense of self-worth and to act autonomously. It is to fail to treat them as free and equal persons.

But what does all this have to do with criminal punishment and abrasive blame? Just like anyone else, a person who has committed a criminal offense (and/or a moral wrong) can and should maintain decent self-respect. Her criminal actions do not somehow carve out her worth as a person; “people are more than the worst thing they have ever done.”²⁸⁹ She cannot draw her sense of worth from misplaced pride in having committed wrongs; doing so would not result in a *morally decent* sense of self-worth. She can, however, draw a morally decent sense of self-worth from other aspects of her personhood. The political community should do what it can to provide her, like anyone else, with the social conditions to restore, develop, or maintain her self-respect.

Many penal institutions, however, do not provide a social basis for self-respect for criminal offenders. On the contrary, penal institutions are often a social environment that threatens the self-respect of criminal offenders. At least as it is practiced in most modern

²⁸⁸ Here I am departing from the letter of Rawls’ work. Rawls does not explicitly argue that political communities should provide a social-basis of self-respect for all members of the community. Others, however, have interpreted his work and applied it to reach this conclusion (or something similar to it). See Shue, “Liberty and Self-Respect,” and Shelby, “Justice, Self-Respect, and the Culture of Poverty.”

²⁸⁹ Helen Prejean, *Dead Man Walking*.

political communities, criminal punishment involves the deprivation of liberty. Liberty is one of the social conditions that is usually necessary for people to have self-respect. If a person is denied liberty to walk around freely, to make daily life choices, or to communicate with others, it is easy (indeed normal) for her to think that she has less worth as a person. But criminal offenders are different from non-offenders in this respect. I agree with philosophers who argue that deprivation of liberty (and perhaps other forms of criminal punishment) – if it is done properly – can express the fact that offenders have equal (rather than inferior) standing as persons in the moral community.²⁹⁰ German penal institutions, at least as they are described by Whitman, stand as examples.²⁹¹

In many penal institutions (especially those in America), however, offenders are not punished in a way that is conducive to their maintenance or recovery of self-respect. The expression of abrasive blame in prisons poses a particularly powerful threat to the self-respect of many criminal offenders. The expressions of abrasive blame, particularly when coupled with the deprivation of liberty, do not help offenders regain or maintain a sense of their intrinsic self-worth or their equal membership in the moral community. Abrasive blame (even the non-degrading abrasive blame that is my focus), in so many words and deeds, tells the offenders that they are trash and scum, *refuse* of the moral community that is begrudgingly kept alive, not full members of it. Even if these

²⁹⁰ See Herbert Morris, “Persons and Punishment”

²⁹¹ If the reader disagrees, then she can take my argument as suggesting a reason to not only refrain from abrasive blame, but also to refrain from depriving criminal offenders of liberty.

expressions are not degrading, they do tend to jeopardize offenders' social basis for self-respect.

Abrasive blame in the criminal justice system tends to jeopardize offenders' social basis for self-respect, but it does not necessarily expunge it. For some offenders, penal conditions that involve abrasive blame may provide an adequate social basis for self-respect. What counts as an adequate social basis will vary based on specific prison conditions and the psychology of individual offenders. Consider the different effects abrasive blame may have on four kinds of offenders. Some fortunate offenders will retain an adequate social basis of self-respect regardless of whether the political community abrasively blames them.²⁹² Offenders in the direst cases will lack an adequate basis of self-respect regardless of whether the political community abrasively blames them or not.²⁹³ Still other offenders' chances of recovering or retaining an

²⁹² For example, consider an offender who has a healthy sense of self-respect before her crime and who receives ample support from an unconditionally loving family or friends after her crime. She does not take any pride in her crime and holds herself accountable for it. But her surroundings do not keep her paralyzed by hopelessness, guilt, or shame; she has the support to move forward and develop the positive aspects of her character. As an example, consider George Bailey at the very end of *It's a Wonderful Life*. Imagine that his friends had come up ten dollars short and he had been arrested and convicted of fraud. He still would have had ample social bases of self-respect regardless of whether the *political* community were to abrasively blame him. The abundance of his social bases of self-respect is (part of) what made him realize that he really did have a wonderful life after all.

²⁹³ For example, consider the former SS officer who is convicted of war crimes in the movie, *The Reader*. Whether or not the state abrasively blames her and sentences her to prison, she lacks adequate resources in society to develop decent self-respect. Her social bases for self-respect do not hang in the balance; they are not fragile, but broken without reasonable hope for adequate repair. A miracle might restore her self-respect, but the state's decision to abrasively blame her or not would just be a drop in this dark well. This case explains why this undermining condition generally does not apply to extremely heinous and publically memorable crimes. Like it or not, such crimes generate enduring social stigmas that tend to effectively destroy adequate social bases for self-respect.

adequate social basis of self-respect would be improved if the political community were to abrasively blame them.²⁹⁴ On the other hand, some offenders' chances of recovering or retaining an adequate basis of self-respect would be significantly diminished if the political community were to abrasively blame them. If the political community were to refrain from abrasively blaming these offenders, their chance of recovering or retaining an adequate social basis of self-respect would be significantly improved.

I call this last type of offenders "fragile offenders." A fragile offender is a criminal offender whose social basis for self-respect "hangs in the balance." I offer the following examples of fragile offenders. They are hypothetical examples based on features of actual offenders. The examples are representative of two ways in which an offender may be fragile: by being deprived of an adequate social basis for self-respect prior to their crime and by committing a crime that makes their self-respect especially vulnerable for the foreseeable future.

Socially Oppressed Offender: Maria is a Hispanic woman who was born in America to parents who had immigrated illegally. At the age of 2, she is an orphan and grows up under the care of the state, bouncing from foster home to foster home because no family wants her. She goes to school where her teachers tell her that she, and other kids "like her" are "slow." She has a relatively low I.Q., she does not speak much English, and kids in school constantly subject her to ethnic slurs and bully her. Foster parents, teachers,

²⁹⁴ Perhaps the abrasive blame will shake the offender of any pre-existing *indecent* self-respect and make room for decent self-respect to develop and grow. Or perhaps the offender has no indecent self-respect but simply has a deep value commitment to the DCAB norm; she believes that "serving her time" or "taking her punishment" is necessary in order to return to society as a person with decent self-respect or honor. Raskolnikov in Dostoyevsky's *Crime and Punishment* is arguably an example of such an offender.

and people in her neighborhood subject her to degrading sexual harassment that she takes to be normal treatment for a woman. She is raped and impregnated at the age of 18.

Maria is both slow and obese. By the time she figures out that she is pregnant, it would be too late to have a legal abortion (moreover, having an abortion at any time is a stigma in her neighborhood). So she has the baby on the street and leaves it by a dumpster. She is convicted of manslaughter. When people ask why she did it, she says that she is trash and what comes out of her is probably trash too. Maria has been deprived of an adequate social basis for self-respect throughout her life and is devoid of self-respect when she commits her crime (she may also be a battered offender). The criminal justice system can respond in a way that plants a seed for self-respect or it can bury any reasonable chance of her having self-respect by incarcerating her in a prison replete with abrasive blame.

Oppressed-by-Guilt Offender: Every summer a handful of parents accidentally leave babies and toddlers in their cars and return to find their children dead, baked and suffocated in the car. There is generally a frantic call to 911 followed by genuine expressions of massive amounts of guilt from the parents. One father was in a catatonic state for several days after killing his son.²⁹⁵ Some of these parents are guilty of culpable violation of criminal law and they are genuinely fragile offenders. They have self-respect prior to the crime, but they lose self-respect upon commission of their crime. If they are subject to the usual abrasive blame offenders receive in prison, most of them will not have an adequate social basis from which to recover their lost self-respect. But if they are spared from such abrasive blame, then they might have an adequate social basis for

²⁹⁵ Weingarten, "Fatal Distraction."

self-respect. Refraining from abrasive blame increases their chance of having a social basis of self-respect.

(I should note that the fragility of these offenders may not undermine the reason to follow the DCAB. It may be that on a careful understanding of American social rules about abrasive blame, abrasive blame toward these parents is inappropriate (presumably because they are “already suffering enough”).)²⁹⁶

When a fragile offender stands before the criminal justice system, the practiced social rules of the community deem it appropriate for the system (as an organ of the political community) to abrasively blame the offender. But, I argue, it would be unfair for the political community to follow these rules and abrasively blame the fragile offender. As I argued above, the political community should, within reason, help provide each of its members with a social basis of self-respect. The political community can, within reason, achieve the justifiable aims of punishment while helping to provide the fragile offender a social basis for self-respect. The political community can do so by sparing the fragile offender from abrasive blame (either by sending her to a prison without abrasive blame, or, if none are available, by giving her a sentence that involves little or no prison time). For example, a judge could sentence a fragile offender like Maria to a halfway house that deprives her of liberty as punishment for her crime, but supports rather than further obstructs her development of self-respect.

²⁹⁶ This is unlikely. Public reactions are often marked by severe abrasive blame. A common type of comment after the acquittal of a mother in this kind of case is: “[the acquitted parent] should have kept her legs closed and not had any kids. They should lock her in a car during a hot day and see what happens,” (Weingarten, “Fatal Distraction”).

A political community that abrasively blames a fragile offender not only misses an opportunity to help provide a social basis for self-respect, but it also puts a social basis for self-respect even further out of reach for the fragile offender. The abrasive blame of the political community actively hinders and obstructs the fragile offender's social basis of self-respect. It is bad enough that a fragile offender thinks of herself as trash. The political community should not fertilize the social soil in which such thoughts take root – but its abrasive blame does just that. To treat a human being as the free and equal person she is, the political community should do what it can to provide her with a social basis of self-respect. By obstructing the fragile offender's social basis of self-respect, the political community fails to treat her as the free and equal person she is. Instead, it treats her as the trash she thinks she is. Assuming that fairness requires treating people as free and equal persons, abrasively blaming fragile offenders is not fair.

Section III: Wounds, not Permissible Stings

There is something unfair (or if one prefers, “otherwise wrong”) about the political community abrasively blaming battered and fragile offenders. But does this mean that it is generally impermissible for the criminal justice system to abrasively blame these offenders? Perhaps not. As explained in the previous chapter, the criminal justice system has a presumptive reason to abrasively blame offenders in accord with the DCAB, the set of decent community-wide social rules about when, how, and to what degree abrasive blame is appropriate. The DCAB in America (and probably other political communities) deems it appropriate for the criminal justice system to abrasively blame battered and fragile offenders (with the possible exception of the guilt-ridden fragile

offender).²⁹⁷ The social rules carve out no recognized exceptions for these offenders. So although there may be something unfair about blaming battered and fragile offenders, there also appears to be a presumptive reason to abrasively blame them.

In this Section, I argue that the unfairness of abrasively blaming battered and fragile offenders washes out the presumptive reason to abrasively blame them. Unless there are other strong considerations in favor of abrasively blaming these offenders, criminal justice systems should refrain from abrasively blaming them and work to carve out a recognized exception for them into the political community's DCAB. The argument takes two different forms.

The first form of the argument insists that social rules which license the criminal justice system to abrasively blame battered and fragile offenders are unjust. Unlike most of the other rules in the political community's DCAB (its entire set of rules about abrasive blame), the rules that deem abrasive blame appropriate for battered and fragile offenders are incompatible with the principles and values in a full theory of political justice. Particularly, the rules seem incompatible with the commitment to treat individuals as free and equal persons. By following the rules and abrasively blaming a battered offender, the political community fails to take seriously her standing as an *equal* member of the community. The rules here fail to instantiate the value of reciprocity. By

²⁹⁷ In general the DCAB tends to deem abrasive blame appropriate (and sometimes required) for all criminal offenders unless their culpability is somehow diminished (they are found not morally responsible or are otherwise excused). Battered and fragile offenders, as I described them, do not have diminished culpability. The DCAB may take other factors into account, such as remorse, but these generally only decrease the appropriate amount of abrasive blame rather than render abrasive blame categorically inappropriate. Moreover, these other factors often do not apply to battered and fragile offenders (they are often lacking in remorse with the exception of the guilt-ridden fragile offender).

following the rules and abrasively blaming a fragile offender, the political community fails to respect the offender's capacity for autonomy by further jeopardizing rather than forwarding her development of that capacity. The rules as applied to a battered or fragile offender fail to take seriously her standing as a being capable of making free and meaningful choices. Even if following these unjust rules is conducive to stability in the community, this does not give the political community good reason to follow them. As discussed in the previous chapter, there is no good moral reason to pursue stability in a way that rests on injustice. Officials in the criminal justice system should take a general stance against abrasively blaming battered and fragile offenders and should work on changing their social rules to make them decent in this respect.

One might argue, however, that the above argument is too quick. I have shown that abrasively blaming battered and fragile offenders is unjust, but I have not shown that the rules allowing such behavior are also unjust. Perhaps there are other reasons standing in favor of the rules which make them compatible with a full theory of justice. If so the rules are decent. Even if this is the case, I argue that the presumptive reason to follow these rules with respect to battered and fragile offenders is outweighed.

The second form of the argument presses not on the decency of the relevant social rules, but on the relative strength of the presumptive reason to follow these rules. As I argued in the previous chapter, the reason to follow the DCAB is presumptive; it generally and non-decisively counts in favor of following the DCAB. The reason does not count in favor of following the DCAB in every instance; it only counts in favor of following the DCAB frequently enough to sustain *general* compliance. Although the political community has reason to maintain general compliance with the DCAB,

refraining from abrasive blame toward battered and fragile offenders usually does not threaten general compliance with the DCAB (for one thing, the social rules usually do not require abrasive blame for battered and fragile offenders, they only deem it appropriate).

So why not refrain from blame when engaging in it would be unfair? The presumptive reason counting in favor of abrasive blame is washed out by the strong consideration of unfairness counting against it. Unless I am missing some general reason counting in favor of abrasively blaming battered and fragile offenders, there is usually no reason that justifies abrasively blaming them. It would not be a permissible sting, but an unfair strike which is marked by a distinct lack of reciprocity and which leaves substantial wounds in its wake.

I am not only arguing that criminal justice officials should refrain (or at least have very strong reason to refrain) from abrasively blaming battered and fragile offenders in individual cases. I am also arguing that they and others outside the system should work to change the rules so as to deem it inappropriate for the criminal justice system to abrasively blame battered and fragile offenders. Even if the existing rules allowing such blame turn out not to be unjust, they are at best minimally decent and we can make them much better, much closer to achieving justice as fairness, by carving out a recognized exception for battered and abused offenders.

It is important to remember that community-wide social rules about abrasive blame are not written in stone. The DCAB is the political community's way to actively work out its reasonable disagreements on abrasive blame, not to have them settled forevermore. The rules tend to be roughly hewn, it is not as though all members of the political community sat down and carefully wrote out the rules with special attention to

capture all the relevant exceptions. As history shows, community-wide social rules about abrasive blame are dynamic. They are refined and changed over time as people stop using certain practices (like branding and the stocks) and start favoring others. The rules change when people change their tune in ways that turn out to command general compliance.²⁹⁸ I have argued that, out of fairness to some of the least advantaged in their community, people should change their tune. Criminal justice systems should ebb the tide and refrain from abrasively blaming these offenders. I hope that individuals' proposed changes will command general compliance; that they will not just permissibly veer off from the herd's well-worn path, but get the herd to change its path. With time and attention, these individuals may carve out a recognized exemption for battered and fragile offenders.

American traditions have given us a criminal justice system that is long on abrasive blame and short on fairness to the least advantaged. James Whitman, a legal scholar whose critique of the American criminal justice system is loosely similar to my own, writes that the harshness of the American criminal justice system is a product of long tradition. Changing this harshness would require "real change...not just in punishment practices but in much grander American cultural traditions. It would be

²⁹⁸ It is important to note that not all changes are permissible. Unlike decreasing the amount of abrasive blame, drastically increasing the amount of abrasive blame is generally not permissible – because one cannot reasonably expect others to accept a justification for doing so (increases have more hurts to justify, not less). Gradually nudging up the amount of abrasive blame while being sensitive to any objections, however, can be permissible. Indeed this is the strategy that the American criminal justice system (including Sheriff Joe Arpaio) has taken.

foolish to think that such change is coming soon.”²⁹⁹ So what is to be done in the meantime? In the conclusion that follows, I argue that positive mercy may not solve our problems, but it might offer us one way to take a step forward.

²⁹⁹ *Harsh Justice*, 207. Whitman identifies degradation (very loosely construed) as the primary problem in American criminal justice, but we have the same general concerns in mind.

CONCLUSION: MERCY AS A BANDAGE?

What is to be done for battered and fragile offenders?

Over the past three chapters, I have investigated abrasive blame and raised concerns about its expression in the criminal justice system. I dispelled a common view that abrasive blame is justified simply in virtue of the fact that a person has committed a wrong action. I argue that abrasive blame is justified under certain conditions and the conditions vary depending on context. In the context of a political community, abrasive blame is justified when it is in accord with decent social rules and the presumptive reason to follow those rules is not outweighed. Given that the justification of abrasive blame is fragile at best, I encourage a full investigation of abrasive blame practices in criminal justice systems to determine their moral permissibility. For now, however, I raise two general concerns about abrasive blame in criminal justice systems. I argue that it would be unfair for political communities to abrasively blame battered offenders and fragile offenders. I also argue that this consideration washes out the presumptive reason to engage in abrasive blame. I suggest that the social rules that call for abrasive blame toward these offenders are potentially unjust. These rules should be reformed so as to recognize exceptions for battered and fragile offenders with respect to abrasive blame in the criminal justice system.

My analysis of abrasive blame does not entail a specific recommendation about how criminal justice systems should punish (or otherwise respond) to battered and fragile

offenders. I recommend future research into whether the arguments I gave against abrasively blaming battered and fragile offenders can be applied to argue for mitigating criminal punishment of these offenders. I am highly sympathetic to such a proposal, but I have not been able to consider it within the limits of this project. Here I emphasize that although abrasive blame is not an essential part of criminal punishment, it is often robustly entwined with the way modern political communities do punish criminals.³⁰⁰ Questions about abrasive blame thereby bear heavily on our decisions about how and even whether to punish criminals in our system.

In some cases, concerns about abrasive blame and punishment conflict and force us to make difficult decisions and compromises. Consider a case in which a) giving a battered or fragile offender Punishment X is the only way to achieve some end required by criminal justice (for example, to send a message that the crime was wrong, to vindicate a system of rights, or to maintain social security) and b) there is no way to give the offender Punishment X without subjecting the offender to abrasive blame (because the wardens keep up a robust practice of abrasive blame toward all offenders in prison). It may be that these conditions are never jointly satisfied in a given political community, but I suspect this may be the case in America. In such a case, the right decision for a judge (prosecutor, police officer, or parole board) might be to give the offender Punishment X. If so, the abrasive blame accompanying that sentence is a necessary evil. Alternatively, the right decision might be to give the offender an alternative that does not

³⁰⁰ Abrasive blame and punishment may not only be entwined in our prisons, but also in the way our brains work. Research suggests that the inclination to blame an offender is a driving force behind most people's judgments of whether to punish the offender. See Lerner, Goldberg, and Tetlock, "Sober Second Thought;" Small and Loewenstein, "The Devil you Know."

involve abrasive blame but which falls short of required ends of corrective justice. If so, falling short of an end required by justice is a necessary evil. Adjudicating between these two alternatives is beyond the scope of my project. What I want to point out, however, is that neither alternative is fully just (even if it is all-things-considered the right decision in a given case). In these conditions, a fully just outcome is out of reach. Second best (or less unjust) is our only option.

A political community need not, however, resign itself to second-best treatment of battered and fragile offenders in the long run. Not only can it work to eliminate (or limit) the incidence of battered or fragile offenders, but it can also work to develop alternative punishments that accomplish the required ends of corrective justice without abrasive blame (for all offenders or just for battered and fragile offenders). The example of German penal institutions suggests that such alternatives are available. With such alternatives in place, the criminal justice system could respond to battered and fragile offenders with a fully just punishment.³⁰¹ The criminal justice system could punish them in a way that would meet required ends of criminal justice, but which would not involve (or be otherwise accompanied by) abrasive blame.

What about Mercy?

In Part I, I identified and defended the exercise of positive mercy in the criminal justice system. The exercise of positive mercy is treating someone justly when unjust

³⁰¹ There are two ways this might work: 1) all offenders, including battered and fragile offenders, are punished without abrasive blame (this is only possible if it would not undermine general compliance with community-wide social rules about abrasive blame) or 2) battered and fragile offenders are punished without abrasive blame but other offenders receive punishment that involves the amount abrasive blame deemed appropriate by the DCAB.

social rules call for a harsher response. If conditions are such that the criminal justice system cannot respond to battered and fragile offenders in a fully just fashion, then the system cannot respond with “full” positive mercy either. The system might rightly spare an offender a harsh treatment that rules would have called for, but the response would not be a perfect instance of positive mercy because the response is less than fully just. Let us call this kind of rightful sparing imperfect positive mercy. (This rightful sparing may be very close to negative mercy – except it is unlikely to be supererogatory.)

Suppose, however, that there are legally available punishments that meet required ends of corrective justice, but which do not involve abrasive blame.³⁰² Officials should derogate from social rules that call for abrasive blame and give battered and fragile offenders these alternative punishments. In doing so, these officials might be exercising law-abiding positive mercy. If social rules about abrasive blame are indeed unjust, then these officials are exercising positive mercy. If the social rules about abrasive blame are minimally decent, then these officials who derogate from the rules and withhold abrasive blame are not exercising positive mercy. Rather, they are exercising something more akin to equity (generally equity is used for behavior that derogates from law, here officials are derogating from a presumably non-legal rule). Or to put it more plainly, the officials are acting rightly by making an exception to a general rule. Although I maintain a theoretical distinction between positive mercy and equity, the distinction in this context does not really matter. What matters in these cases is that the officials refrain from

³⁰² Working to create alternative punishments that do not involve abrasive blame may be a kind of proto-positive mercy. It is work that we need to do in order to practice positive mercy.

abrasive blame and treats battered and fragile offenders justly. If even a few officials act in this way, then we will already be one small step forward.

Cases concerning abrasive blame may often (and likely do) arise that give officials occasion for exercising positive mercy that derogates from law. It is in these cases that the two parts of my dissertation best work together to shed light on Lenin's question. Suppose (not unrealistically) that an unjust³⁰³ mandatory minimum sentencing law requires a judge to give a battered or fragile offender a sentence that is longer than what is needed to serve justified ends of criminal punishment. To make things both easy and realistic, let us say the sentence is twenty-five years for petty theft. I think the judge in this case has a reason to exercise positive mercy regardless of whether the offender is battered or fragile. If the offender is battered or fragile, however, the judge has an even better case to make for positive mercy. Sentencing the offender for twenty-five years would not only incapacitate him (the "need" to incapacitate recidivist offenders is generally cited as the reason for three-strikes laws), but it would also subject him to

³⁰³ I stipulate that the law in this case is unjust. If in other cases the law is not unjust, then the judge has occasion to exercise not positive mercy, but equity. I am not entitled to conclude that the judge has standing reason to exercise equity in such cases. Although I am highly sympathetic to the proposal that judges would have reason to exercise equity in such cases, I have not argued for this proposal in the dissertation. Other legal philosophers argue that judges have reason to exercise equity and I am sympathetic to their arguments. I have not, however, defended these (or other) arguments in this dissertation. I gave a defense of positive mercy, not equity, in Chapter 2. My *suspicion* is that this defense can be applied to equity without considerable difficulty. Equity – at least when it is limited to the context of criminal punishment – seems easier to defend than positive mercy. Equity stays within the spirit of the laws and is thereby less (if only slightly less) subject to objections that it violates the judicial obligation to uphold the rule of law. Given that mercy meets these objections, I suspect that equity with respect to criminal punishment does too. But I cannot be sure. Equity might pose unique challenges that I have not explored here. I leave a full investigation of equity for future work.

twenty-five years of near continuous abrasive blame given the current state of American prisons. The judge can argue that the political community should not sting the offender with abrasively blame at all, let alone do so for twenty-five years. This is not merely an impermissible sting; it is constant and wrongful *abuse*.

If the judge orders such abuse on the grounds that she has “no choice” but to apply the law, her order inflicts a deep wound on both the offender and the rule of law in the political community. This is a prime occasion for the exercise of positive mercy. My work in the first part of the dissertation does more than give a label to the action of a judge who makes the brave decision to derogate from the unjust law and give the offender just treatment in the form of an alternative sentence (in the absence of an available alternative to blame-ridden incarceration, the judge can at least shorten the sentence to months). In Chapter 2, I argued that the judge who makes this decision stands up not only for the offender, but also for the rule of law. She need not weigh whether to uphold the law or do substantive justice to the offender, she can stand firm in the belief that she is doing her job: pursuing justice under the rule of law.³⁰⁴

³⁰⁴ This argument is a new way in which a “rotten social background” can mitigate sentencing – namely by rendering abrasive blame unfair. My approach differs from those who suggest a “rotten social background” can mitigate punishment only by excusing offenses of the least advantaged on grounds that they are somehow less responsible for their actions. This latter approach has been objected to on grounds that it vitiates the autonomy of the least advantaged. The “rotten social background” excuse as it has been interpreted thus far in the literature, is said to insult the least advantaged by denying them full moral agency and responsibility for their actions. See Stephen Morse, “Deprivation and Desert,” as well as his four-part debate with Judge Bazelon in *Southern California Law Review*. The argument I have offered, however, might be applied to mitigate punishment without any implication that offenders have diminished responsibility or culpability.

Her action is not negative mercy; she is giving the offender what he does deserve in the form of a just alternative sentence, she is moved by reasons and a firm sense of justice rather than pity or charity, and her action is not optional from the moral point of view. She may well do the action without a warm, willing heart or any special concern for the particularities of the offender. She may be angry at both the state and the offender for putting her in a position where she has to stick her own neck out to do the right thing. I mention this not because it would be a good thing if the judge felt this way, but to highlight the contrast that I developed in Chapter 1 between negative mercy and positive mercy as it is exhibited in this case.³⁰⁵

One might wonder why I insist on calling positive mercy “mercy” instead of simply justice. (Why not leave mercy to the defenders of negative mercy?) Although positive mercy is a part of justice, there is merit in giving positive mercy its own name and account as I did in Chapter 1. First, I am interested in the idea of making mercy work in a criminal court. I have argued that positive mercy has a great deal of potential

³⁰⁵ I have no objection to those who continue to work on defending negative mercy. I simply want to show that positive mercy has different and significant work to do in criminal justice systems. Although my project has concentrated on positive mercy, the second part of my project may be of use to those interested in negative mercy. Writers about negative mercy, such as Tasioulas (in “Mercy”) and Murphy (in *Forgiveness and Mercy*) and others, claim that an offender who has suffered “positive evil” or has crippling remorse should be granted negative mercy. It is left unclear exactly *why* these facts should mitigate what these authors take to be deserved punishment. My work in the second part of the dissertation might be applied to explain what is going on here. Many offenders who have experienced positive evil or crippling remorse are battered or fragile offenders. I have explained why these offenders should be given less abrasive blame than the amount generally correlated with their culpability. This work can be applied to explain why – in some cases – these offenders should be given less punishment (generally punishment involves abrasive blame to which they should not be subject) than they deserve (deserved punishment is generally taken to be the amount correlated with culpability).

for use in this context. But in the context of the modern courtroom, the term “justice” often refers to a very specific and narrow kind of justice: justice that is in accord with valid laws (“legal justice”). Even though the judicial exercise of positive mercy serves the rule of law, it is a departure from particular laws and so it departs from this kind of justice. It may, therefore, be misleading to conflate mercy with justice in the context of the courtroom.

Second, when we talk about “justice” in the courtroom, we generally assume that the judge has a right to exercise it. Similar to a citizen who engages in a morally right act of civil disobedience, however, the judge who exercises positive mercy acts rightly – but without a right to so act. I follow Raz in thinking that there are two ways one can show that one is entitled to some action. One can show that one’s action is right or one can show that one has a right to the action; “to show that one has a right to it is to show that even if it wrong, one is entitled to perform it.”³⁰⁶ The judge who shows positive mercy must demonstrate that her action is right; her case rests on its own moral merit. If “positive mercy” is wrong, she is not entitled to show it. In this sense, mercy is an *exceptional political action*. Mercy is part of justice, but it is an exceptional part and giving it a name of its own helps us to appreciate this fact.

Third, and finally, mercy is almost always understood as a response to imperfection in both human beings and, as I argue here, in human institutions. Mercy reminds us that there are always problems, even ones that we do not yet know about, that we have to work on upon our muddy planet. In this, it keeps us humble. Moreover, mercy keeps us active. It does not purport to solve the imperfections to which it

³⁰⁶ Raz, *The Authority of Law*, 274.

responds. It is not an achievement that has any laurels upon which to rest. Justice keeps us asking what we should be aiming for, but the part of justice that is mercy keeps us asking: What is to be done *now*?

REFERENCES

- Abbott, Jack Henry. *In the Belly of the Beast*. New York: Random House, 1981.
- Adler, Jacob. "Murphy and Mercy." *Analysis* 50, no. 4 (1990): 262-268.
- Anselm. *Proslogion with the replies of Gaunilo and Anselm*. Translated by Thomas Williams. Indianapolis: Hackett, 2001.
- Aristotle. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett, 1999.
- Aronson, Ronald. "Responsibility and Complicity." *Philosophical Papers* 19, no. 1 (1990): 53-73.
- Barber, Nick. "Must Legalistic Conceptions of the Rule of Law Have a Social Dimension?" *Ratio Juris* 17 (2004): 474-488.
- Barrozo, Paulo D. "Punishing Cruelly: Punishment, Cruelty, and Mercy." *Criminal Law and Philosophy* 2, no. 1 (2008): 67-84.
- Bazelon, David, L. "The Morality of the Criminal Law." *Southern California Law Review* 49 (1976): 385-405.
- . "The Morality of the Criminal Law: A Rejoinder to Professor Morse." *Southern California Law Review* 49 (1976): 1269-1274.
- Bazelon, Emily. "Arguing Three Strikes." *The New York Times Magazine*, May 23, 2010: MM40.
- Beck, A. "The Rule of Law and Legal Rules." *Mind* 82, no. 325 (1973): 104-105.
- Beccaria, Cesare. *An Essay on Crimes and Punishments: the original fourth edition*. Edited by Adolph Caso. Boston: International Pocket Library, 1983.
- Bennett, Christopher. "The Limits of Mercy." *Ratio* 17 (2004): 1-11.
- Bentham, Jeremy. *Of the Limits of the Penal Branch of Jurisprudence*. In *Collected Works of Jeremy Bentham*, edited by P. Schofield. Oxford: Oxford University Press, 2010.
- . *Principles of Penal Law*. In *The Works of Jeremy Bentham, Volume 1*, edited by John Bowring. Edinburgh: William Tait, 1843.
- . *The Constitutional Code*. In *The Works of Jeremy Bentham, Volume 9*, edited by John Bowring. Edinburgh: William Tait, 1843.

- Berlin, Isaiah. "Two Concepts of Liberty." In Berlin, *Four Essays on Liberty*. Oxford: Oxford University Press, 1969.
- Brady, Karren. "Porridge that's Hard to Stomach." *Birmingham Evening Mail*, June 17, 2008: News 10.
- Braithwaite, John and Philip Pettit. *Not Just Deserts: A Republican Theory of Criminal Justice*. New York: Oxford University Press, 1990.
- Brandt, R. B. "Conscience (Rule) Utilitarianism and the Criminal Law." *Law and Philosophy* 14, no. 1, Conrad Johnson Memorial Issue (1995): 65-89.
- Braund, Susanna Morton, ed. and trans. *De Clementia*. Oxford: Oxford University Press, 2009.
- Brett, Nathan. "Mercy and Criminal Justice: A Plea for Mercy." *The Canadian Journal of Law and Jurisprudence* 5 (1992): 347-360.
- Brien, Andrew. "Mercy and Desert." *Philosophical Papers* 20 (1990): 193-201.
- . "Mercy, Utilitarianism, and Retributivism." *Philosophia* 24 (1995): 493-521.
- . "Mercy within Legal Justice." *Social Theory and Practice* 24, no. 1 (1998): 83-110.
- . "Saving Grace: A Reply to Murphy." *Criminal Justice Ethics* 9 (1990): 52-59.
- Brown, Ian "It Took a Jester to Point the Finger." *The Globe and Mail*, March 14, 2009.
- Card, Claudia. "On Mercy." *The Philosophical Review* 81, no.2 (1972): 182-207.
- Cartwright, David. "Revenge, Punishment, and Mercy: The Self-Overcoming of Justice." *International Studies in Philosophy* 17 (1985): 17-26.
- Coleman, K. M. "Fatal Charades: Roman Executions Staged as Mythological Enactments." *The Journal of Roman Studies* 80 (1990): 44-73.
- "Con or Pawn?" *Ottawa Citizen*, July 20, 2000: A14.
- Corlett, J. Angelo. "Making Sense of Retributivism." *Philosophy* 76 (2001): 77-110.
- Dan-Cohen, Meir. *Harmful Thoughts: Essays on Law, Self, and Morality*. Princeton: Princeton University Press, 2002.
- Davis, Michael. "Punishment as Language: Misleading Analogy for Desert Theorists." *Law and Philosophy* 10, no. 3 (1991): 311-322.

---. "Sentencing: Must Justice be Even-Handed?" *Law and Philosophy* 1, no. 1 (1982): 77-117.

Dolinko, David. "Mismeasuring 'Unfair Advantage': A Response to Michael Davis." *Law and Philosophy* 13, no. 4 (1994): 493-524.

---. "Morris on Paternalism and Punishment." *Law and Philosophy* 18, no. 4 (1999): 345-361.

---. "Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment." *Law and Philosophy* 16, no. 5 (1997): 507-528.

---. "Some Thoughts about Retributivism." *Ethics* 101, no. 3 (1991): 537-559.

---. "Some Naive Thoughts about Justice and Mercy." *Ohio State Journal of Criminal Law* 4 (2007): 349-360.

Doris, John M. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press, 2002.

Dostoyevsky, Fyodor. *Crime and Punishment*. Translated by David McDuff. London: Penguin, 1996.

---. *The Brothers Karamazov*. Translated by David McDuff. London: Penguin, 2003.

Duff, Antony. *Answering for Crime: Responsibility and Liability in the Criminal Law*. Oxford: Hart, 2007.

---. "The Intrusion of Mercy." *Ohio State Journal of Criminal Law* 4 (2007): 361-387.

Dworkin, Ronald. *A Matter of Principle*. Cambridge: Harvard University Press, 1985.

---. "Law's Ambitions for Itself." *Virginia Law Review* 71, no. 2 (1985): 173-187.

---. *Law's Empire*. Cambridge: Harvard University Press, 1986.

---. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge: Harvard University Press, 2000.

---. *Taking Rights Seriously*. Cambridge: Harvard University Press, 1978.

Dylan, Bob. "Masters of War." In *Freewheelin'*. Sony, 1990 compact disc. Originally released in 1963.

Dyzenhaus, David. *Hard Cases in Wicked Legal Systems: Pathologies of Legality*. Oxford: Oxford University Press, 2010.

- . *Legality and Legitimacy: Carl Schmitt, Hans Kelsen and Hermann Heller in Weimar*. Oxford: Oxford University Press, 1997.
- , ed. *Recrafting the Rule of Law: The Limits of Legal Order*. Portland, OR: Hart, 1999.
- Egmond, Florike. "The Cock, the Dog, the Serpent, and the Monkey. Reception and Transmission of a Roman Punishment, Or Historiography as History." *International Journal of the Classical Tradition* 2, no. 2 (1995): 159-192.
- Ehlers, S., V. Schiraldi, and J. Ziedenberg. "Still striking out: Ten years of California's three strikes." Washington, DC: The Justice Policy Institute. June 2004. <http://www.justicepolicy.org/article.php?list=type&type=76>.
- Feinberg, Joel. *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton: Princeton University Press, 1970.
- . "The Expressive Function of Punishment." *The Monist* 49 (1965): 397-423.
- . *The Moral Limits of the Criminal Law*. New York: Oxford University Press, 1984-1988.
- . "The Nature and Value of Rights." *Journal of Value Inquiry* 4 (1970): 243-257.
- Feldman, Fred. "Desert: Reconsideration of Some Received Wisdom." *Mind* 104, no. 413 (1995): 63-77.
- Ferrajoli, Luigi, and Danilo Zolo. "Marxism and the Criminal Question." *Law and Philosophy* 4, no. 1 (1985): 71-99.
- Finnis, John. "Helping Enact Unjust Laws without Complicity in Injustice." *American Journal of Jurisprudence* 49 (2004): 11-42.
- . *Natural Law and Natural Rights*. Oxford: Oxford University Press, 1980.
- . "On Reason and Authority in *Law's Empire*." *Law and Philosophy* 6, no.3 (1987): 357-383.
- Fuller, Lon L. *The Morality of Law*. New Haven: Yale University Press, 1969.
- Garcia, J. L. A. "Two Concepts of Desert." *Law and Philosophy* 5, no. 2 (1986): 219-235.
- Gardner, John B. "Complicity and Causality." *Criminal Law and Philosophy* 1 (2007): 127-141.

- . *Offences and Defences: Selected Essays in the Philosophy of Criminal Law*. New York: Oxford University Press, 2007.
- . "Review: Complicity: Ethics and Law for a Collective Age." *Ethics* 11, no. 4 (2004): 827-830.
- . "The Legality of Law." *Ratio Juris* 17 (2004): 168-181.
- Garland, David. *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago: University of Chicago Press, 2001.
- Garvey, Stephen. "'As the Gentle Rain From Heaven': Mercy in Capital Sentencing." *Cornell Law Review* 81 (1996): 989-1048.
- Geach, Peter. *Truth and Hope: The Furst Franz Josef Und Furstin Gina Lectures Delivered at the International Academy of Philosophy in the Principality of Liechtenstein*. Notre Dame: University of Notre Dame Press, 2001.
- "Glass of Milk 'Saved' Nazi Victim." *One-Minute World News*. BBC News, January 27, 2006. http://news.bbc.co.uk/2/hi/uk_news/wales/46.
- Good Will Hunting*. DVD. Directed by Gus Van Sant. New York: Miramax, 1997.
- Gross, Hyman, and Ross Harrison, eds. *Jurisprudence: Cambridge Essays*. New York: Oxford University Press, 1992.
- Gur-Ayre, Miriam. "Comment [on A Theory of Complicity by Sanford Kadish]." In *Issues in Contemporary Legal Philosophy: The Influence of H.L.A. Hart*, edited by Ruth Gavison, 304-310. Oxford: Clarendon Press, 1987.
- Hampton, Jean. "Democracy and the Rule of Law." *Nomos* 36 (1994): 13-44.
- Harrison, Ross. "The Equality of Mercy." In *Jurisprudence: Cambridge Essays*, edited by Hyman Gross and Ross Harrison. Oxford: Clarendon Press, 1992.
- Hart, H. L. A. *Punishment and Responsibility: Essays in the Philosophy of Law*. New York: Oxford University Press, 2008.
- . *The Concept of Law*. 2nd ed. Oxford: Clarendon, 1994.
- Hart, H. L. A., and Lon L. Fuller. "The Morality of Law." In *Essays in Jurisprudence and Philosophy*, edited by H. L. A. Hart. Oxford: Clarendon, 1983.
- Hasebe, Yasuo. "The Rule of Law and its Predicament." *Ratio Juris* 17 (2004): 489-500.

- Hassine, Victor. *Life Without Parole: Living in Prison Today 4th Edition*. Edited by Robert Johnson and Ania Dobrzanska. New York: Oxford University Press, 2008.
- Hayek, F. A. *The Road to Serfdom*. London: Routledge, 1976.
- Heffernan, William C. and John Kleinig, eds. *From Social Justice to Criminal Justice: Poverty and the Administration of Criminal Law*. New York: Oxford University Press, 2000.
- Hestevold, H. Scott. "Justice to Mercy." *Philosophy and Phenomenological Research* 46, no. 2 (1985): 281-291.
- Hill, Thomas E. Jr. "Symbolic Protest and Calculated Silence." *Philosophy and Public Affairs* 9, no. 1 (1979): 83-102.
- . "Treating Criminals as Ends in Themselves." *The Annual Review of Law and Ethics* 11 (2003): 17-36.
- Holmes, Oliver Wendell, Jr. "The Path of the Law." In *The Path of the Law and its Influence: The Legacy of Oliver Wendell Holmes, Jr.*, edited by Steven J. Burton. New York: Cambridge University Press, 2000.
- Hurd, Heidi M. "The Morality of Mercy." *Ohio State Journal of Criminal Law* 4 (2007): 389-407.
- Husak, Douglas. "Holistic Retributivism." *California Law Review* 88, no. 3 (2000): 991-1000.
- . "On the Supposed Priority of Justification to Excuse." *Law and Philosophy* 24, no. 6 (2005): 557-94.
- . "Why Punish the Deserving?" *Noûs* 26, no. 4 (1992): 447-64.
- It's a Wonderful Life*. DVD. Directed by Frank Capra. Los Angeles: Republic Entertainment, 1947.
- Johnson, Carla Ann Hage. "Entitled to Clemency: Mercy in the Criminal Law." *Law and Philosophy* 10, no. 1 (1991): 109-118.
- Kadish, S. "Complicity, Cause and Blame: A Study in the Interpretation of Doctrine." *California Law Review* 73 (1985): 323-410.
- Kadish, Mortimer R. and Sanford H. Kadish. *Discretion to Disobey: A Study of Lawful Departures from Legal Rules*. Stanford: Stanford University Press, 1973.

- Kant, Immanuel. *The Metaphysics of Morals*. Translated and edited by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.
- Kennedy, Anthony M., Associate Justice, Supreme Court of the United States. Address at the American Bar Association Annual Meeting. August 9, 2003. http://www.supremecourt.gov/publicinfo/speeches/viewspeeches.aspx?Filename=sp_08-09-03.html.
- King, Martin Luther, Jr. *Letter from a Birmingham Jail*. New York: Harper Collins, 1994.
- Kleinig, John. "Mercy and Justice." *Philosophy* 44 (1969): 341-342.
- . "The Concept of Desert." *American Philosophical Quarterly* 8, no. 1 (1971): 71-78.
- Kobil, Daniel T. "Should Mercy Have a Place in Clemency Decisions?" In *Forgiveness, Mercy and Clemency*, edited by Austin Sarat and Nasser Hussain. Stanford: Stanford University Press, 2007.
- Kramer, Matthew H. "Requirements, Reasons, and Raz: Legal Positivism and Legal Duties." *Ethics* 109, no.2 (1999): 375-407.
- Kutz, Christopher. *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press, 2000.
- Lerner, J., J. Goldberg, and P. Tetlock. "Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility." *Personality and Social Psychology Bulletin* 24 (1998): 563-574.
- Lyons, David. "Justification and Judicial Responsibility." *California Law Review* 72, no. 2 (1984): 178-99.
- Markel, Dan. "Against Mercy." *Minnesota Law Review* 88 (2004): 1421-1480.
- Marmor, Andrei. "The Rule of Law and its Limits." *Law and Philosophy* 23, no.1 (2004): 1-43.
- May, Larry. *Sharing Responsibility*. Chicago: Chicago University Press, 1992.
- May, Larry and Stacey Hoffman, eds. *Collective Responsibility*. Maryland: Rowman and Littlefield, 1991.
- Mill, John Stuart. *Collected Works of John Stuart Mill*. Edited by J. M. Robson. Toronto: Toronto University Press, 1963.

- Mills, Eugene. "Scheffler on Rawls, Justice, and Desert." *Law and Philosophy* 23, no. 3 (2004): 261-72.
- Moore, Kathleen Dean. *Pardons: Justice, Mercy, and the Public Interest*. New York: Oxford University Press, 1989.
- Moore, Michael, S. *Placing Blame: A General Theory of the Criminal Law*. New York: Oxford University Press, 1997.
- . "The Moral Worth of Retribution." In *Character, Responsibility, and the Emotions*, edited by F. Schoeman. Cambridge: Cambridge University Press, 1987.
- More, Thomas. *Utopia*. Edited and translated by Paul Turner. London: Penguin Books, 2003.
- Morris, Herbert. *On Guilt and Innocence: Essays in Legal Philosophy and Moral Psychology*. Berkeley: University of California Press, 1976.
- . "Persons and Punishment." *The Monist* 52 (1968): 475-501.
- Morse, Stephen J. "Deprivation and Desert." In *From Social Justice to Criminal Justice*, edited by W. Hefferman and J. Kleinig, 114-160. New York: Oxford University Press, 2000.
- . "The Twilight of Welfare Criminology: A Reply to Judge Bazelon." *Southern California Law Review*, 49 (1976): 1247-1268.
- . "The Twilight of Welfare Criminology: A Final Word." *Southern California Law Review*, 49 (1976): 1275-1276.
- Murphy, Colleen. "Lon Fuller and the Moral Value of the Rule of Law." *Law and Philosophy* 24, no. 3 (2005): 239-262.
- Murphy, Jeffrie G. "Marxism and Retribution." *Philosophy and Public Affairs* 2, no. 3 (1973): 217-243.
- . "Violence and the Rule of Law." *Ethics* 80, no.4 (1970): 319-21.
- Murphy, Jeffrie G., and Jean Hampton. *Forgiveness and Mercy*. Cambridge: Cambridge University Press, 1988.
- Nagel, Thomas. *Equality and Partiality*. Oxford: Oxford University Press, 1991.
- Nietzsche, Friedrich. *On the Genealogy of Morality*. Edited by Keith Ansell-Pearson. Translated by Carol Diethe. Cambridge: Cambridge University Press, 1994.

- Norrie, Alan. *Punishment, Responsibility, and Justice*. Oxford: Oxford University Press, 2000.
- Nussbaum, Martha C. "Equity and Mercy." *Philosophy and Public Affairs* 22, no. 2 (1993): 83-125.
- Nuyen, A. T. "Straining the Quality of Mercy." *Philosophical Papers* 23 (1994): 61-74.
- O'Driscoll, Lyla H. "The Quality of Mercy." *Southern Journal of Philosophy* 21 (1983): 229-250.
- Oliver, William. "The Streets: An Alternative Black Male Socialization Institution." *Journal of Black Studies* 36, no. 6 (2006): 918-937.
- Olsaretti, Serena, ed. *Desert and Justice*. New York: Oxford University Press, 2003.
- Pater, Joshua. "Struck out Looking: Continued Confusion in Eighth Amendment Proportionality Review after *Ewing v. California*." *Harvard Journal of Law and Public Policy* 27 (2003): 399-421.
- Pearn, John. "The Quiddity of Mercy – A Response." *Philosophy* 71 (1996): 603-604.
- Pojman, Louis. "Equality and Desert." *Philosophy* 72 (1997): 549-570.
- Pojman, Louis P., and Owen McLeod, eds. *What Do We Deserve?: A Reader on Justice and Desert*. New York: Oxford University Press, 1999.
- Postema, Gerald J. "Custom in International Law: A Normative Practice Account." In *The Nature of Customary Law: Legal, Historical, and Philosophical Perspectives*, edited by J. B. Murphy and A. Perreau-Saussine. Cambridge: Cambridge University Press, 2007.
- . "Law's Ethos: Reflections on a Public Practice of Illegality." *Boston University Law Review* 90 (Symposium 2010): 1847-1868.
- Prejean, Helen. *Dead Man Walking*. New York: Vintage Books, 1994.
- Quint, David. *Montaigne and the Quality of Mercy: Ethical and Political Themes in the Essais*. Princeton: Princeton University Press, 1998.
- Rachel Getting Married*. DVD. Directed by Jonathon Demme. New York: Clinica Estetico, 2008.
- Radin, Margaret Jane. "Reconsidering the Rule of Law." *Boston University Law Review* 69 (1989): 781-819.

- Radin, Max. "The Lex Pompeia and the Poena Cullei." *The Journal of Roman Studies* 10 (1920): 119-130.
- Rainbolt, George W. "Mercy: An Independent, Imperfect Virtue." *American Philosophical Quarterly* 27 (1991): 169-173.
- . "Mercy: In Defense of Caprice." *Noûs* 31, no. 2 (1997): 226-241.
- Ransome, W. "'Above the Sceptred Sway': Retrieving the Quality of Mercy." *Critica* 40 (2008): 3-27.
- Rawls, John. *A Theory of Justice*. Cambridge: Belknap Press of Harvard University Press, 1999.
- . *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge: Belknap Press of Harvard University Press, 2001.
- . *Political Liberalism*. New York: Columbia University Press, 1993.
- Raz, Joseph. *The Authority of Law*. Oxford: Oxford University Press, 1983.
- . *Ethics in the Public Domain*. Oxford: Oxford University Press, 1996.
- Rebel Without a Cause*. DVD. Directed by Nicholas Ray. Burbank, CA: Warner Bros. Pictures, 1955.
- Reinhardt, Tobias. "Introduction." In *Dialogues and Essays*. Edited and translated by John Davie. New York: Oxford University Press, 2009.
- Rideau, Wilbert. *In the Place of Justice: A Story of Punishment and Deliverance*. New York: Alfred A. Knopf, 2010.
- Ripstein, Arthur. "Equality, Luck, and Responsibility." *Philosophy and Public Affairs* 23, no. 1 (1994): 3-23.
- . *Equality, Responsibility, and the Law*. Cambridge: Cambridge University Press, 1999.
- Roberts, H. R. T. "Mercy." *Philosophy* 46 (1971): 352-3.
- Roberts, Julian V. "Public Opinion and Mandatory Sentencing: A Review of International Findings." *Criminal Justice and Behavior* 30, no. 4 (2003): 483-508.
- Sadurski, Wojciech. "Distributive Justice and the Theory of Punishment." *Oxford Journal of Legal Studies* 5, no. 1 (1985): 47-59.

- Sarat, Austin and Nasser Hussain, eds. *Forgiveness, Mercy, and Clemency*. Stanford: Stanford University Press, 2007.
- Scanlon, T. M. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge: Belknap Press of Harvard University Press, 2008.
- . *What We Owe to Each Other*. Cambridge: Belknap Press of Harvard University Press, 1998.
- Schapiro, Tamar. "Compliance, Complicity, and the Nature of Nonideal Conditions." *Journal of Philosophy* 100, no. 7 (2003): 329-355.
- Scheffler, Samuel. "Doing and Allowing." *Ethics* 114, no. 2 (2004): 215-39.
- . "Justice and Desert in Liberal Theory." *California Law Review* 88, no. 3, The Morality of Criminal Law: A Symposium in Honor of Professor Sandy Kadish (2000): 965-990.
- . "Relationships and Responsibilities." *Philosophy and Public Affairs* 26, no. 3 (1997): 189-209.
- . "Responsibility, Reactive Attitudes, and Liberalism in Philosophy and Politics." *Philosophy and Public Affairs* 21, no. 4 (1992): 299-323.
- Sebba, Leslie. "The Pardoning Power: A World Survey." *The Journal of Criminal Law and Criminology* 68, no. 1 (1977): 83-121.
- Seneca, Lucius Annaeus. "De Clementia." In *De Clementia*, edited and translated by Susanna Morton Braund. Oxford: Oxford University Press, 2009.
- . "On Clemency." In *Anger, Mercy, Revenge*, edited and translated by Robert A. Kaster and Martha C. Nussbaum. Chicago: The University of Chicago Press, 2010.
- . "On Mercy." In *Moral Essays, Volume I*, edited and translated by John W. Basore. London: Loeb Classical Library, 1928.
- . "On Mercy." In *Dialogues and Essays*, edited and translated by John Davie. New York: Oxford University Press, 2009.
- . "On Anger." In *Dialogues and Essays*, edited and translated by John Davie. New York: Oxford University Press, 2009.
- Shakespeare, William. *The Merchant of Venice*. Edited by Sylvan Barnet. New York: Signet Classics, 2005.

- Shakespeare, William. *Henry VI*. Edited by Barbara A. Mowat and Paul Werstine. Folger Shakespeare Library. New York: Washington Square Press, 1992.
- Shelby, Tommie. "Justice, Self-Respect, and the Culture of Poverty." Paper presented at Philosophy Department Speaker Series, University of North Carolina, Chapel Hill, NC, January 15, 2010.
- Sher, George. *Desert*. Princeton: Princeton University Press, 1987.
- . *In Praise of Blame*. Oxford: Oxford University Press, 2006.
- Shklar, Judith. "Political Theory and the Rule of Law." In *Political Thought and Political Thinkers*, edited by Stanley Hoffman. Chicago: University of Chicago Press, 1998.
- Shue, Henry. "Liberty and Self-Respect." *Ethics* 85, no. 3 (1975): 195-203.
- Sigler, Mary. "The Story of Justice: Retribution, Mercy and the Role of Emotions in the Capital Sentencing Process." *Law and Philosophy* 19 no. 3 (2000): 339-367.
- Small, D., and G. Loewenstein, "The Devil You Know: The Effects of Identifiability on Punitiveness." *Journal of Behavioral Decision Making* 18, no. 5 (2005): 311-318
- Smart, Alwynne. "Mercy." *Philosophy* 43 (1968): 345-359.
- Smiley, Marion. *Moral Responsibility and the Boundaries of Community*. Chicago: Chicago University Press, 1992.
- Statman, Daniel. "Doing without Mercy." *The Southern Journal of Philosophy* 32 (1994): 331-354.
- Sterba, James. "Can A Person Deserve Mercy?" *Journal of Social Philosophy* 10 (1979): 11-14.
- Strange, Carolyn, ed. *Qualities of Mercy: Justice, Punishment, and Discretion*. Vancouver: UBC Press, 1996.
- Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75 (1994): 5-24.
- Strawson, Peter. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1962): 1-25.

- Steiker, Carol. "Tempering or Tampering? Mercy and the Administration of Criminal Justice." In *Forgiveness, Mercy and Clemency*, edited by Austin Sarat and Nasser Hussain. Stanford: Stanford University Press, 2007.
- Stuntz, William J. "The Pathological Politics of Criminal Law." *Michigan Law Review* 100 (2001): 505-598.
- Tasioulas, John. "Mercy." *Proceedings of the Aristotelian Society* 103 (2003): 101-132.
- The Reader*. DVD. Directed by Stephen Daldry. Los Angeles: Mirage Enterprises, 2008.
- Twambley, P. "Mercy and Forgiveness." *Analysis* 36, no. 2 (1976): 84-90.
- U.S. Bureau of Justice Statistics. "Correctional Populations in the United States, 2009." <http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=2316>.
- U.S. Census Bureau. "State and County QuickFacts." August 2010. <http://quickfacts.census.gov/qfd/index.html>.
- Valcke, Catherine. "Civil Disobedience and the Rule of Law. A Lockean Insight." *Nomos* 36 (1994): 45-62.
- Vinciguerra, Thomas. "The Clothes that Make the Inmate." *The New York Times*. October 1, 2000: 4-2.
- Von Hirsch, Andrew, Andrew Ashworth, and Julian Roberts, eds. *Principled Sentencing: Readings on Theory and Policy*. Portland, OR: Hart, 2009.
- Waldron, Jeremy. "Is the Rule of Law an Essentially Contested Concept (in Florida)?" *Law and Philosophy* 21, no. 2 (2002): 137-164.
- . "Rights in Conflict." *Ethics* 99, no. 3 (1989): 503-519.
- . *The Dignity of Legislation*. Cambridge: Cambridge University Press, 1999.
- . "The Role of Rights in Practical Reasoning: 'Rights' versus 'Needs.'" *The Journal of Ethics* 4, no. 1 (2000): 115-135.
- . "The Rule of Law in Contemporary Liberal Theory." *Ratio Juris* 2 (1989): 79-96.
- . "Why Law – Efficacy, Freedom, or Fidelity?" *Law and Philosophy* 13, no. 3 (1994): 259-284.
- Walker, Nigel. "The Quiddity of Mercy." *Philosophy* 70 (1995): 27-37.

- Wallace, R. Jay. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments." In *Reasons and Recognition. Essays on the Philosophy of T. M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. New York: Oxford University Press, forthcoming.
- Walmsley, Roy. "World Prison Population List (Eighth Edition)." King's College London International Centre for Prison Studies (2008).
www.kcl.ac.uk/depsta/law/research/icps/.../wppl-8th_41.pdf.
- Watson, Gary. "Responsibility and the Limits of Evil." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by F. Schoeman, 256-286. New York: Cambridge University Press, 1987.
- Weingarten, Gene. "Fatal Distraction: Forgetting a Child in the Backseat of a Car is a Horrifying Mistake. Is it a Crime?" *Washington Post*, March 8, 2009.
- Weisberg, Robert. "Reappraising Complicity." *Buffalo Criminal Law Review* 4 (2000): 217-281.
- Whitman, James Q. *Harsh Justice: Criminal Punishment and the Widening Divide between America and Europe*. New York: Oxford University Press, 2003.
- Wolf, Susan R. "Blame Italian Style." In *Reasons and Recognition. Essays on the Philosophy of T. M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. New York: Oxford University Press, forthcoming.
- . *Freedom within Reason*. New York: Oxford University Press, 1990.
- . "The Moral of Moral Luck." In *Setting the Moral Compass: Essays by Women Philosophers*, edited by Cheshire Calhoun. Oxford: Oxford University Press, 2004.
- Yeager, Daniel. "Helping, Doing, and the Grammar of Complicity." *Criminal Justice Ethics* 15, no.1 (1996): 25-35.
- Young, Neil. "Ambulance Blues." In *On the Beach*. Reprise/Wea, 2003 compact disc. Originally released in 1974.
- Zaitchik, Alan. "On Deserving to Deserve." *Philosophy and Public Affairs* 6, no. 4 (1977): 370-388.