

Introduction

But how canst thou spare the wicked if thou art wholly just and supremely just? For how does the wholly and supremely just do something that is not just? . . . We can find no reason to explain why, among men who are equally evil, thou does save some, and not others, through thy supreme goodness, and does condemn the latter, and not the former, through thy supreme justice.¹

The problem of “justifying” mercy is old,² but it has resurfaced recently in light of truth and reconciliation commissions and other transitional justice methods, debates over discretion in the federal sentencing guidelines, debates over the executive’s clemency powers,³ and debates over “restorative justice” alternatives to traditional, state-imposed retributive punishment.⁴ We wonder whether a reconciliation that does not involve punishment can be just; we wonder whether an executive is ever right to pardon out of compassion; we wonder whether judges should have discretion to sentence leniently in cases where defendants are remorseful, evoke compassion, have dependents, are ill, or are community heroes.

As the preceding quotation illustrates, St. Anselm articulated the problem as well as any contemporary writer: if God is all-just, then how can he be at the same time merciful? For if he is merciful, then he treats like cases unlike and is thereby unjust. Mercy, for Anselm, is the grace that lightens or eliminates deserved punishment. Justice is treating like alike, according to desert. To be both just and merciful, then, would seem to be logically impossible.

Within the retributive punishment tradition that understands punishment as some form of “just deserts,” nearly all of the current philosophical literature on mercy either seeks to discredit mercy’s place in public punishment or seeks to find a limited place for it “to the side” of regular forms of retribution.⁵ The

most common arguments of this kind are (1) that mercy has a place only as a form of equity, when needed to give retributive principles more play than is possible within the strict legal rules;⁶ (2) that mercy as forgiveness has a place in “private” relationships but not usually in the context of public punishment—except perhaps in very limited circumstances in which victims have given the government the “go ahead” to forgive;⁷ (3) that mercy may have a limited role of reestablishing peace in cases where justice is impossible or unworkable;⁸ or (4) that mercy as compassion may be useful as an emotional state to better discern justice (determining, rather than reducing, just punishment).⁹

Efforts to advocate theories of punishment that give pride of place to mercy as forgiveness and reconciliation¹⁰ have not yet been given a complete conceptual and philosophical grounding within mainstream philosophical traditions.¹¹ They are instead often advocated on policy or utilitarian grounds as the best we can do within the realm of the possible, given the practical limitations of *real* justice.¹² Mercy as an act of reducing or eliminating just punishment, in other words, must be (at most) an exceptional, not central, part of the legal system.

The main hurdle for reconciliatory theories, as it has been since Anselm, is the threefold retributivist objection that mercy is demeaning, inegalitarian, and unjustifiable. As Justice Clarence Thomas said, “A system that does not hold individuals accountable for their harmful acts treats them as less than full citizens. In such a world, people are reduced to the status of children or, even worse, treated as though they are animals without a soul.”¹³ This book tries to dig out the roots of retributivism to address these objections.

A different way of eliminating the mercy problem—one that I do not discuss in this book—is to reject retributivism’s premises and to define punishment in utilitarian terms, by arguing that we should exact whatever penalty best reduces overall suffering from criminal activity. Desert, in such a model, is just a side constraint or no constraint at all: if the greater suffering caused by crime could be eliminated by the lesser suffering of punishing a few innocent people, well, why not do this, except to the extent that it might create more suffering in the long term by undermining public confidence in the penal system? Lenity, in utilitarian terms, would be just another value factored into an analysis of social benefits: if a lenient sentence prevents more crime or causes less suffering overall than a heavier sentence, then lenity is appropriate. However, utilitarian theories of punishment have various difficulties,¹⁴ the keenest of which is a variant of Justice Thomas’s Kantian objection: utilitarian theories treat defendants as objects to be manipulated by incentives and penalties, rather than persons capable of making moral choices.

If one takes the Kantian view that defendants ought not be treated as objects to be manipulated by sticks and carrots but must be understood as persons making autonomous moral choices and deserving punishment when they make selfish ones, employing lenity as a “carrot” seems demeaning and wrong. We should not do right selfishly, to get a benefit; doing right is itself a requirement of unselfishness. For these reasons, this book treats mercy as a problem within a framework of retributive punishment and says little about utilitarian theories. The challenge is to confront retributivism’s Kantian account of moral personhood directly.

Hence, to defend mercy requires an argument about why reason pace Kant is not the ground of responsibility, ethics, and community. I will argue that reason cannot be such a ground because reason itself requires a prior stance of being with others. In doing so, I flip the accepted relationship between justice and mercy on its head. Instead of mercy being the antithesis or exception to law or, at best, a practical compromise, mercy takes over as the ground of justice, the basis on which justice itself is possible. The philosophical understanding of moral personhood switches from a kind of simplified Kant that has become legal catechism—what I call *kanticism*—which purports to ground current retributive approaches to sentencing, to an understanding derived from the later Kant of the *Critique of Judgment*, Martin Heidegger, and Emmanuel Levinas. Switching our philosophical lens shows a very different account of the relationship between justice and grace and, derivatively, between punishment and mercy.

The goal of this book is to let mercy appear in a different light and to speak to both the analytic and continental philosophical traditions from a starting place laid out by their common ancestor, Kant. Part 1, comprised of chapters 1–4, undertakes the task of seeing mercy and punishment—and, consequently, law—from a different philosophical perspective.

Chapter 1 explains the problem caused by our reliance on kanticism. The central problem is that, in our kanticism, we believe that reason comes *before* our connection to each other, as a necessary “glue” linking person to person, a condition of the possibility of ethical community. Because mercy does not follow rules of reason and cannot be universalized, we believe it to be in conflict with the possibility of community, destabilizing our reason-mediated relations with each other and denying us our place within a “kingdom of ends.” Yet reason itself, thought as a system of perfectly articulated universal rules, is not within our grasp, and we are left chasing shadows, unable to cross the unfathomable distance between one person and another.

Chapter 2 moves to a new philosophical starting place and introduces Heidegger and Levinas. These thinkers switch the problem around: ethical community is the given, and reason is derivative. We do not need reason as the glue of community; we are already *with* others. Mercy, as a compassionate gift to an undeserving guilty person, no longer is destabilizing or contrary to human nature but reflects the human situation: we are at core those who receive the gift of world and others without deserving that gift. We do not *earn* community through choosing to be reasonable; it is already given. Law changes, too. It is understood no longer as universal rules of reason but as a tradition that guides us by analogy, not deduction. Law is no longer syllogism but common law—a provisionally stated “holding” that can be refined and revised as the rule hits the road in future cases. The ability to *see* the “distinction with a difference” between cases is perceptive, not deductive—a way of already being in tune with others and with one’s world. As with community, the perception of likeness or difference that guides legal analysis comes *before* rules and enables us to articulate them. This “givenness” is at the heart of both mercy and law: it is a “gift” that we neither earned nor made nor deserved but that enables both law and community.

Chapter 3 sketches out how punishment theory is implicated in the problems with kanticism. Retributivism thinks of punishment as restoring us to ethical community by repairing the maxims of our actions to be fully universalized laws, applicable to ourselves as well as others. The retributive story is that, as we experience the rebound of our own imperfect rule, the rule is perfected and we are thereby restored to solidarity with others as reasonable persons. The problem we discover is that there is no human standpoint from which such a universal rule can be established. Retributivism is therefore not a possibility for us.

Chapter 4 is the heart of this book and introduces a theory of punishment and mercy from our different philosophical starting place. Wrong is no longer the irrational but the indifferent, no longer a failure to live up to our nature as rational but a failure to live up to our nature as “with others.” Punishment is no longer the universalizing of a maxim but the pain of a shared remembering of one’s self-alienation from the connection and responsibility for others. The pain of this experience seeks resolution in a settlement or sacrifice that will rekindle trust and ethical community for the future. This “settlement” or possible future is not “deserved” but is the return of our own being-with. In this way, punishment becomes inextricably linked with mercy as an “undeserved” leniency. Punishment culminates in an unsecured, undeserved settlement or

compromise that takes an unguaranteed risk of being-with and proffers trust for a future, rather than settling accounts with the past. Punishment is the restoration of being-with, rather than the restoration of reason.

In part 2, comprising chapters 5 and 6 and the conclusion, I play out this new view of mercy in more concrete legal contexts of pardon and sentencing, using both legal and narrative examples. My methodology in these chapters matches the philosophical insights gained in part 1. For this reason, I do not try to construct *rules* of mercy that I then apply to various contexts. Instead, consistent with the idea that law is *given* by tradition and perception in a way that is temporally extended, incomplete, and based in common-law reasoning, I try to tease an account of mercy out of the concrete contexts already provided in the legal and cultural materials I study. I examine these “cases” to try to learn something about how mercy *already is at work* in our legal system. Viewed through the lens of being-with, we perceive our legal system differently; some practices that we usually consider exceptional come to seem central, and others that seemed central move to the background. Philosophical study, like the common law, becomes a matter of reading these past cases and articulating provisional “holdings” that partially illuminate a way forward, by seeing patterns and family resemblances in what is already there in our tradition. Because we are human and finite, we cannot pretend to articulate universal principles that will hold for all time. We must be content to feel our way forward by drawing on the richness of what has come before.

Chapter 5 addresses the question of when pardons might or might not be appropriate and how we can think about these ethical questions from a non-Kantian perspective. Chapter 6 addresses the question of whether there are unforgivable crimes and what mercy might mean in such cases. The conclusion builds on two novels—Alexandre Dumas’ *Count of Monte Cristo* and J. M. Coetzee’s *Disgrace*—that are usually read as triumphs of retributive justice. I read these novels to dramatize the fault lines of retributivism and to reflect on how punishment can be understood anew as a shared memory of wrong that calls for mercy as a nondemeaning, undeserved settlement. Even in these stories, as in the law generally, I find mercy already at work.