THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF PHILOSOPHY

CRIMINAL JUSTICE AS FAIRNESS: A RAWLSIAN APPROACH TO JUSTIFYING CRIMINAL PUNISHMENT

BRENDAN BERNICKER
SPRING 2019

A thesis
submitted in partial fulfillment of the requirements
for baccalaureate degrees
in Philosophy and Political Science
with honors in Philosophy

Reviewed and approved* by the following:

Jonathan Marks
Associate Professor of Bioethics, Humanities, and Law
Thesis Supervisor

Brady Bowman
Associate Professor of Philosophy
Honors Advisor

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Mass incarceration raises fundamental questions about how a political society ought to regard those who break its laws and the limits that ought to exist on the power of the state to impose suffering and deprivation on individuals or groups. These questions address issues at the heart of political philosophy, yet that discipline has remained nearly silent on criminal punishment over the past half-century. In this paper, I offer a justification for a practice of criminal punishment that is grounded in a Rawlsian public political conception of justice. On this view, punishment exists to rehabilitate offenders and protect society. Furthermore, all of the suffering that results from the imposition of punishment must be considered in establishing a practice of punishment—including both the suffering that is experienced by offenders' loved ones or communities and that which is imposed on offenders extrajudicially, such as through sexual or other violence during incarceration. Where this suffering can be neither justified nor eliminated, it cannot be imposed. While the argument that I offer applies to the Rawlsian well-ordered society (with its assumption of full compliance relaxed), I use the last section of the paper to discuss some ways in which the current practice of criminal punishment in the United States falls far short of the theoretical ideal and results in unjustifiable suffering and deprivation.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First and foremost, I want to thank my family for their unconditional love and support. I have no idea where I would be without them, but I am certainly glad that I do not have to find out. I also want to thank my friends for being the best proofreaders, bandmates, dining companions, unlicensed therapists, and late-night interlocutors that I could have asked for.

More directly, I am grateful to Prof. Jonathan Marks for always making time in his busy schedule to give me advice and constructive criticism on this paper, and for holding me to a high standard throughout the thesis-writing process. I also appreciate his insistence that I never lose sight of how this theoretical project relates to the real people who are adversely affected by unjust systems of punishment.

This project grew out of two classes that I took with Profs. Chris Moore and John Christman, both of whom have been invaluable mentors during my time at Penn State. I am grateful to both of them, as well as to Profs. Amy Linch, Ben Jones, Brady Bowman, Uygar Abaci, and Vincent Colapietro for inspiring and supporting me over the past four years.

I would be remiss if I did not thank the Rock Ethics Institute for its generous support of this project, and for hosting me this past year as a Thesis & Research Fellow. I am also incredibly grateful to Prof. Catherine Wanner and the Paterno Fellows Program for their unwavering confidence in me, and for bending over backward to further my intellectual, professional, and personal development.

Finally, I want to thank Prescott Loveland, and the clients whom I have served as a criminal defense investigator, for inspiring me to pursue a career in public interest law and a research agenda that focuses on making the criminal justice system fairer and more equitable.

Chapter 1

Introduction

For the past half-century, liberal political philosophy has been nearly silent on questions of criminal punishment.[1] This is unfortunate for at least two reasons. First, this time period has seen an unprecedented expansion of incarceration and criminal punishment more generally. Political philosophers have only recently begun to grapple with the ethical consequences of mass-incarceration, and they lack adequate theoretical frameworks for critiquing one of the most pressing social and political issues of our time.

Second, the silence of political philosophers on questions of criminal justice has meant that the philosophy of law has essentially developed parallel to political philosophy. Philosophers who deal with crime and punishment tend to situate their arguments either within theories of the nature and purpose of law itself, or within broader critiques of prevailing social structures of oppression and domination: there have been relatively few systematic attempts to justify institutions of punishment with reference to a broader political conception of justice.

Part of the blame for this rests on John Rawls.[2] As many others have observed, "Rawls set the agenda of political philosophy in the late twentieth century and still sets it now, well into the twenty-first."[3] By almost entirely omitting questions of punishment from his work, Rawls turned the conversations in mainstream political philosophy away from crime and punishment.

---

[1] Chad Flanders, "Criminals behind the Veil: Political Philosophy and Punishment," *Brigham Young University Journal of Public Law* 31, no. 1 (2016): 83–109.

[2] Ibid., 84-85.

[3] Ibid., 84.

The result of this is that "we lack, for the most part, any worked-out theory of what the policing and processing of crime should look like."[4]

Mass incarceration raises fundamental questions about how a political society ought to regard those who break the law and what limits ought to exist on the power of the state to impose suffering and deprivation on individuals or groups. These questions, and particularly the latter of the two, address issues at the heart of political philosophy. They can no longer be ignored.

In this paper, I attempt to use Rawls' theory of Justice as Fairness to establish criteria for evaluating both systems of punishment and particular impositions of punishment. By grounding these criteria in a widely-respected theory from political philosophy, I hope to close the rift between the philosophical justifications for penal institutions and those for other social institutions. I also hope to provide a framework for evaluating punishment with reference to a political, rather than comprehensive, conception of justice.

In the sections that follow, I will introduce the Rawlsian concept of public justification through a political conception of justice, and explain that it pertains to justifying the "basic structure" of society. Subsequently, and after demonstrating that it is appropriate to relax Rawls' assumption of full compliance with respect to criminal behavior, I will apply Rawls' two principles of justice to the practice of punishment. Finally, I will explore the implications of this type of justification for particular impositions of punishment, which are justified by reference to the rules that define the practice of punishment rather than directly by reference to the principles of justice.

---

[4] Ibid., 83.

Chapter 2

Groundwork for a Justification of the Practice of Punishment

Before I can begin my attempt to establish justificatory criteria for punishment, it is important that I am first clear about the type of justifications that I intend to offer. Specifically, I will justify punishment by referring exclusively to a public political conception of justice.

To say that a conception of justice is public is to say that it is recognized by all reasonable members of a political society. Because Rawlsian political societies are assumed to be reasonably plural, their members will not all share comprehensive conceptions of justice, such as those that are informed by particular religious or philosophical traditions.[5] Accordingly, some people may believe that justice requires retributive punishment, where individuals who do wrong things deserve to suffer in proportion to the wrongness of their actions, while others may believe that justice requires all punishment to have a strictly utilitarian function. While individuals are free to hold these and other personal comprehensive conceptions of justice, such conceptions will vary from person to person and therefore can never be public in the sense required to justify a social institution.

---

[5] Importantly, reasonable pluralism allows for a wide range of comprehensive conceptions of the good and their accompanying comprehensive conceptions of justice. It does not, however, allow for all comprehensive conceptions of the good. Specifically, a well-ordered society does not tolerate conceptions of the good or of justice that are incompatible with the public political conception of justice that regulates the political society. All members of a society must recognize the basic liberties of every other member, and comprehensive conceptions of the good that require that non-adherents be deprived of their basic liberties, including their equal political liberties, are not tolerated.

Recognizing this, Rawls offers a political conception of justice. This conception is not based on any particular religious or philosophical tradition, and does not assume that citizens share a conception of the good.[6] Instead, it explicitly assumes that citizens will have different or even competing conceptions of the good, and attempts to set out the basic structure of a fair system of cooperation that regards individuals as free and equal citizens who are reasonable and rational.[7] The political conception of justice also regards citizens as capable of developing "the two moral powers," which are (1) "the capacity to understand, apply, and act from (rather than merely in accordance with) the principles of political justice" and (2) "the capacity to have, revise, and rationally pursue a conception of the good."[8]

Deriving and defending this conception of political justice is the primary purpose of Rawls' *Justice as Fairness*, and I will not repeat those arguments here. This is not, and does not purport to be, a comprehensive introduction to Rawls' political philosophy generally or even to Justice as Fairness in particular. Instead, I will introduce the elements of Rawls' political conception of justice as they become relevant to understanding punishment in a Rawlsian well-ordered political society (in which crime occurs for reasons other than traditional non-compliance).

---

[6] A conception of the good is a "family of final ends and aims which specifies a person's conception of what is of value in human life or, alternatively, what is regarded as a fully worthwhile life." These conceptions are generally related to comprehensive religious, philosophical, or moral doctrines in that the aims and ends that compose a conception of the good are ordered and understood through such a doctrine, and the elements of a conception of the good are usually "set within, and interpreted by," comprehensive doctrines. *See* John Rawls, *Justice as Fairness: A Restatement*, (Cambridge, MA: The Belknap Press of Harvard University Press, 2001), 19.

[7] In this paper, I will refer to the members of a political society as its "citizens." I do not use this term to refer exclusively to those people who can claim legal citizenship in the society.

[8] Rawls, *Justice as Fairness*, 18-9.

First, Rawls' political conception of justice takes the basic structure of society as its object. The basic structure is "the way in which the main political and social institutions of society fit together into one system of cooperation." Because the criminal law and practices of punishment are quintessential parts of the basic structure, it is perfectly appropriate to apply Rawls' two principles of justice to them. These principles are that:

> 1. Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and
>
> 2. Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (the difference principle).[9]

While it may be tempting to refer directly to these principles in evaluating a particular imposition of punishment, it is important to remember that they apply to the "political and social institutions" of society, not necessarily to particular acts that fall under those institutions.[10] To understand this distinction, it is helpful to consider another of Rawls' most influential works.

---

[9] Ibid., 42-3.

[10] Throughout this paper, I will refer to practices of punishment and institutions of punishment interchangeably. I do not use those terms in the same sense as MacIntyre, though much of what he says of practices is true also of practices in the Rawlsian sense. The aim of this project is to justify a practice of punishment within the framework of Justice as Fairness, but I welcome efforts to explore the purpose of, and constraints on, punishment in the context of other traditions in political philosophy. It strikes me that evaluating whether justice is a MacIntyrian internal good in the context of an appropriate practice of criminal law and punishment may be a particularly fruitful avenue for future research in this field, though I am not familiar enough with MacIntyre's work to carry out such a project.

On the Two Types of Justification for Punishment

In "Two Concepts of Rules," Rawls explores the importance of distinguishing between the justification for the institution of punishment (what Hart calls the "general justifying aim" of punishment) and the justifications for particular impositions of punishment within the larger institution.[11] The difference, as Rawls explains, can be illustrated by considering two questions a child might ask of a parent: 'why do we put people in jail?' (about the general justifying aim) and 'why was person X put in jail yesterday?' (about a particular imposition).[12] As Rawls notes, we may very well offer justifications at either level that are inapplicable at the other level.[13]

The practical implication of this is that, while legislators may enact statutes in order to promote particular goods, judges generally enforce statutes without reference to the goods that justified their enactment.[14] If, for example, a legislature were to institute an outright ban on smoking cigarettes on the grounds that cigarette smoke adversely affects public health, it would not be a defense to prosecution for a particular smoker to claim, even truthfully, that their smoking did not adversely affect public health. The rationale that supports the prohibition in general does not necessarily also support each particular instance of the prohibition's

---

[11] John Rawls, "Two Concepts of Rules," *The Philosophical Review* 64, no. 1 (January 1955): 3.

[12] Ibid.

[13] Specifically, Rawls argues that utilitarian considerations justify the existence of a system of punishment, while retributive considerations justify particular impositions of punishment. I will argue that Justice as Fairness can offer justifications distinct to each level but compatible with one another and superior to those that Rawls used to illustrate his point in 1955.

[14] This is an over-simplification of the relationship between judging and legal rules. It is not uncommon in legal systems for there to be legal rules that require judges to note the purpose for which a statute was enacted and ensure that particular applications of that statue serve that purpose (or at least do not undermine that purpose). The point here is that the justification for an action within a practice of positive law comes from legal rules, not directly from first principles or natural facts about the world.

enforcement: instead, the enforcement of the prohibition is justified by the prohibition's existence as a rule.

On the Practice Conception of Criminal Justice

That latter observation, that the enforcement of the prohibition is justified by the prohibition's existence as a rule, is the essence of the practice conception of rules, which Rawls famously distinguishes from the "summary conception" of rules.[15] In the summary conception, rules summarize how principles were directly applied to particular situations in the past. When similar situations arise in the future, moral agents can consult these summaries to determine what result the direct application of a principle to the circumstances is likely to yield, and rules form to guide conduct as circumstances recur.[16]

Importantly, these rules are not themselves the principles that justify acting in a particular way, as the principles are logically prior to the rules. Instead, rules merely save time and cognitive effort by obviating the need to apply principles directly to circumstances where those circumstances have previously arisen and been satisfactorily addressed.[17] Individuals may decide to ignore a rule and directly apply the underlying principle to their circumstances (as the moral force of the rule comes only from its concordance with the principle); but because the rule is

---

[15] Because I am focusing on Rawls' philosophy in this project, I will not elaborate on the relationship between the practice conception of rules and theories of rule utilitarianism, except to clarify that the practice conception of rules can apparently accommodate justifications that are not in consequentialist terms.

[16] Ibid., 18.

[17] Ibid., 19.

itself a summary of how this application has been resolved in the past, the result is likely to be the same.[18]

In the practice conception of rules, rules define a practice— or a system of defined offices, moves, and offenses.[19] To illustrate how this differs from the summary conception of rules, Rawls invites readers to consider a game of baseball: while somebody can swing a piece of wood or slide onto a bag outside of the context of a game of baseball, they cannot be said to "strike out" or "steal a base" unless they are also playing baseball.[20] The concepts of "striking out" and "stealing bases" are only meaningful because of the rule-like practice that constitutes a game of baseball, and the rules that define a game of baseball are therefore prior to any particular action that is specified by those rules.[21]

While rules are justified by external considerations, they themselves justify the actions they define. Keeping with the baseball example, the MLB rules prohibit balking "to prevent the pitcher from deliberately deceiving the base runner," but a pitcher who drops the ball while touching the rubber has committed a balk even if no baserunners are deceived.[22] That dropping the ball in these circumstances violates a rule of baseball is sufficient reason to call a balk and

---

[18] Ibid.

[19] Ibid., 20-1.

[20] Ibid.

[21] An extended version of this baseball analogy can be found at Rawls, John. "Two Concepts of Rules." *The Philosophical Review* 64, no. 1 (January 1955): 20-21.

[22] Interestingly, and as with some legal prohibitions, MLB rules include a statement that "Umpires should bear in mind that the purpose of the balk rule is to prevent the pitcher from deliberately deceiving the base runner. If there is doubt in the umpire's mind, the 'intent' of the pitcher should govern." The prohibition on dropping the ball, however, explicitly includes unintentional dropping. *See* Official Playing Rules Committee, *Official Baseball Rules: 2018 Edition* (Office of the Commissioner of Baseball, 2018): 72.

advance the baserunners, and no umpire need appeal to the purpose of the rule against balking in order to make the call.

It is common among philosophers of law to regard positive law as this sort of rule-like practice, and I will sketch out the precise nature of this rule-like practice later in the paper.[23] For now, it is sufficient to note that, once a general practice of punishment is justified and in place, impositions of punishment pursuant to that practice will be justified by conforming to the practice's rules rather than by separate moral or prudential principles. The fundamental task of justifying punishment in the Rawlsian well-ordered society, then, is offering a general justifying aim for the practice of punishment.

On Relaxing the Assumption of Full Compliance

Before I can offer an account of this general justifying aim, I must first address two more aspects of Rawls' theory of Justice as Fairness. The first, which is often considered an obstacle to applying Justice as Fairness to the criminal context, is the idea that Rawls is doing ideal theory and assumes citizens' perfect compliance with the demands of political justice. In the Rawlsian society, then, there is no crime to respond to or punish.[24]

---

[23] H. L. A. Hart, *The Concept of Law*, 3rd ed., Clarendon Law Series (Oxford: Oxford Univ. Press, 2012).

[24] There is still crime to prevent. While Rawls never devotees sustained attention to the criminal law in *Justice as Fairness*, he does spend a few pages in *A Theory of Justice* laying out an interesting and relevant argument about the role of threats of punishment in a society where no crime occurs. That argument begins with a simple premise: "While we may know ourselves, or think we do, and know that we are motivated by our sense of justice to obey the law and follow the principles of justice, we do not know that this is the case with others" (Flanders 2016, 90). Because we are unsure of others' motives, we are unwilling to ground the stability of our political society entirely on the belief that others will comply with the demands of justice

If crime does not exist in ideal theory, how might the principles of justice apply to it? Perhaps crime, like reasonable pluralism, can be considered in ideal theory on the grounds that it will always exist in a free society, and Rawlsian ideal theory is only "realistically," not absolutely, utopian. Of reasonable pluralism, Rawls writes:

> There is no politically practicable way to eliminate this diversity except by the oppressive use of state power to establish a particular comprehensive doctrine and to silence dissent, the fact of oppression (§11). This seems evident not only from the history of democratic states but also from the development of thought and culture in the context of free institutions.[25]

Is the answer, then, to say something similar of criminal behavior? Ought we insist that the history of democratic states tells us that there is no way to eradicate crime, except possibly by oppressive state power? In a word, no.

Rawls goes to great lengths in much of his work to explain why reasonable pluralism goes beyond being a brute fact about society, and is instead normatively desirable. Leaving those arguments aside, however, there is a more important distinction: what makes the use of state power to crush pluralism "oppressive," is that it seeks to "establish a particular comprehensive

---

willingly. To borrow some Kantian language, if we cannot ensure that others will act from duty, it is important to society that they at least act in accordance with duty. As a result, "ideal theory requires an account of penal sanctions as a stabilizing device" (Rawls 1999, 212), or a way to "underwrite citizens trust in one another" (Rawls 1999, 505). Because Rawls is clear that he is not talking about actually punishing anyone, as he is still assuming full compliance with the demands of justice, my reading of these sections is that these incentives serve to assure the members of the society that full compliance will persist, rather than to actually ensure that full compliance will persist. *See* John Rawls, *A Theory of Justice: Revised Edition* (Cambridge, MA: The Belknap Press of Harvard University Press, 1999).

[25] Rawls, *Justice as Fairness*, 84.

doctrine" and impose conceptions of the good upon people.[26] If citizens are to be regarded as free and equal, and especially as both reasonable and rational, then their second moral power to form, revise, and pursue conceptions of the good must be respected. The state power is therefore oppressive because it uses political power in service of comprehensive, rather than political, conceptions of justice.

This argument does not apply to criminal punishment, at least where the criminal laws are justified, for the simple reason that the power of the state is being used in service of the political conception of justice. Where an individual's conduct extends beyond those liberties that are compatible with every citizen's enjoyment of a coextensive scheme of liberties, the individual's conduct can be censured and constrained with reference only to the political conception of justice.

The assumption of full compliance cannot be relaxed, then, simply because full compliance has not occurred in any actual society. It should be relaxed, however, because of the reasons why it has never occurred in any actual society. Rawls regards compliance as a moral choice, which is why non-compliance triggers a desert of punishment. This view rests on an unreasonably broad concept of free will.

In *A Theory of Justice*, Rawls considers a hypothetical person who does not experience the affirmation of his sense of justice as a good, and who therefore "chafes against the principles of justice, rather than finding them at least consistent (if not congruent) with [their] good."[27] Of

---

[26] Ibid.; This can be understood as referring to theocracies, ideologically-totalitarian regimes, perfectionist regimes, and similar systems that attempt to eradicate reasonable pluralism as "oppressive."

[27] Flanders, Chad, "Criminals Beyond the Veil," 92.

such a person, Rawls claims that "their nature is their misfortune."[28] Because the principles that underlie the criminal law are collectively rational, the fact that they are not individually rational to our hypothetical individual is of no consequence. Society is entitled to enforce its just laws against the willfully noncompliant.

As I intend to show, however, Rawls' political principles of justice require that the effect of unearned and undeserved advantage or disadvantage on social outcomes be minimized. To the extent that crime is partially the result of "misfortune," of the kind Rawls identified or the more common kind I will turn to shortly, its control is a social responsibility as well as a personal one. The idea that an individual should be expected to bear the entire burden of their "misfortune" is inconsistent with the rest of Rawls' work, and I am not aware of any other place in his major writings in which he makes an analogous claim. To see why this is, it is important to consider one final aspect of Rawls' political theory: the original position.

On the Original Position

To Rawls, the principles that govern the basic structure of society are chosen by representatives from a menu of possible principles of justice. Any potential principles known to any representative must be added to the menu at their request, but the concept of selecting from a menu illustrates that the principles of justice cannot be directly derived from Rawls' background assumptions or the structure of his decision making context, which is known as the original position.

---

[28] Rawls, *A Theory of Justice*, 504.

In the original position, representatives know basic facts about humans and human society, as well as history and the social sciences. These representatives must choose principles, however, from behind a "veil of ignorance" that prevents them from knowing the particular conceptions of the good they will have in the society that they design.[29]

The argument goes that, acting as trustees for the free, equal, reasonable, rational citizens of society, these representatives will select principles that maximize the position of the worst-off member of society (since they or their constituents may be that person when the veil is lifted). The representatives will also attach all opportunities for social advantage or a greater share of society's resources (deviations from equality of outcomes) to offices that are available to all under fair conditions of opportunity. Because of the first constraint, these advantageous offices are only allowed to exist when their primary benefit accrues to the least-advantaged members of society. Taken together, these statements comprise the second principle of justice.

To see what this might look like in practice, imagine a society determining how to compensate doctors. In this society, all doctors must incur significant opportunity costs in order to attend medical school, which is physically and intellectually demanding. Some doctors work on research that benefits everyone in society, others provide direct services to society's least-advantaged members, and some provide direct services exclusively to advantaged members of society. Because health is a primary good, those in poor health are considered less advantaged than those in good health.

---

[29] It is sometimes said that the veil of ignorance prevents representatives from knowing the "places" that their constituents will occupy in society. It is important to remember that the "places" that exist in the society to be occupied are, in a particular sense, consequences of decisions made in the original position.

First, society would allow those doctors who work on basic research or direct services to the least advantaged to receive a greater share of social resources than others— as these disparities would incentivize people to incur the costs and risks of training to become doctors, and would therefore result in an improvement in the position of the least-advantaged members of society. Society would not allocate a greater share of resources to those doctors who exclusively serve society's advantaged members.

Second, because the office of "doctor" carries a disparately high share of resources, at least for those who serve the least advantaged, it must be attainable under fair conditions of opportunity. This means that admission to medical school, and whatever other educational programs are prerequisite to attending medical school, must be contingent only on permissible factors (such as potential to excel as a doctor) and these programs must be affordable to everyone who is admitted. These requirements come directly from the second principle of justice.

The first principle of justice is also selected from the original position: the representatives do not know what conception of the good they (or their constituents) will hold, so they maximize the extent to which all individuals are free to pursue their conceptions of the good.[30] This means agreeing to an efficient scheme of basic liberties where no person can enjoy any more liberty without infringing on the liberties of others.[31]

It is important to note that the original position is merely a representation device: there is no assembly of representatives behind a veil of ignorance, but those interested in determining

---

[30] Relatedly, representatives do not know whether they or their constituents will be in a minority group, so they ensure that members of minority groups enjoy the same liberties as everyone else.

[31] This is also an implication of Kant's moral law, as each person may enjoy only those liberties that can be universalized.

principles of justice ought to reason "in accordance with the modeled constraints, citing only reasons those constraints allow."[32] In addition to following the constraints on information (primarily the veil of ignorance) and on representation (of free and equal citizens possessed of the moral powers), the final relevant constraint on the original position procedure is that all political principles must be public, in the sense discussed previously.

Finally, in discussing social obligations, it is important to bear in mind that the society Rawls is writing about is composed of free and equal members. We all share in the burdens of social obligations as well as in their corresponding entitlements. Analogously, any and all uses of the political power of the state against citizens are also undertaken by their fellow free and equal citizens.[33] In Rawls' society, citizens are both subject and sovereign.

---

[32] Rawls, *Justice as Fairness*, 86.

[33] "Political power is always coercive power applied by the state and its apparatus of enforcement; but in a constitutional regime political power is at the same time the power of free and equal citizens as a collective body. Thus political power is citizens' power, which they impose on themselves and one another as free and equal." Ibid., 40.

Chapter 3

On the General Justifying Aim of A Practice of Punishment

While representatives in the original position must conceive of citizens as reasonable and rational in selecting the principles that govern the basic structure of society, they are free to acknowledge that, under different circumstances and at different times in their lives, many people experience compromised or underdeveloped capacities of reason. Distributing the burdens of diminished reasoning is a task that must be undertaken from the original position.

This is obviously true for children. Very young children cannot conceive of, and be moved to action by, principles of political justice freely selected through the procedure represented as the original position; and those who can select these principles would be aware of this sociological fact. As representatives and trustees, they would select principles of justice that ensure that young children's moral powers are carefully cultivated, so that they can learn to entertain and be moved by concerns of justice, and so that they can form, revise, and pursue conceptions of the good.

Of course, young children will not be the only members of a group who require social support in developing the moral powers: those with psychosocial or neurological disorders may also require support; and from the original position, representatives aware that they or some of their constituents may have such a disorder would ensure that society shares in the burden of this support.

With respect to the first moral power in particular, there are both epistemic and motivational components: someone with an adequately developed first moral power will both know what political justice requires (the epistemic component) and be moved by this knowledge to conform their conduct with the demands of political justice (the motivational component). Where somebody has not yet learned to distinguish right from wrong, they require the *cultivation* of the epistemic component of the first moral power; but where somebody does wrong knowingly, they require the *fortification* of the motivational component of the first moral power. Society has an obligation to support individuals in both *cultivating* and *fortifying* their first moral powers.

This support can take many forms including, for those who commit crimes, rehabilitative intervention. To illustrate this point, I will show how it applies to three groups of offenders, though there may well be others that I do not deal with explicitly. These groups are those who commit crimes because they do not understand that their conduct is wrong, those who understand the demands of justice but experience an inability to conform their conduct to those demands, and those who know that their conduct is wrong and willingly engage in it regardless of the demands of justice.[34]

Start with those who do not recognize that their conduct is wrong (in the public political sense of the term, which generally means unlawful). In this example, the individual has committed an epistemic moral failure. Once we accept, as I demonstrated above, that there is a societal obligation to cultivate one another's first moral power, crimes rooted in ignorance of the demands of justice reflect a social failure as well as an individual one: we are all responsible for

[34] Reconciling the appropriate handling of crimes of necessity with Rawlsian ideal society is beyond the scope of this project.

cultivating our own first moral powers, but we are also responsible for cultivating the first moral powers of our fellow citizens.

While the offending individual must temporarily be prevented from further impinging on the liberties of others, possibly by incarceration, this is a prophylactic rather than punitive measure. The individual is not to be given 'hard treatment,' but is instead to be quarantined, in a sense, while learning what justice requires. The goal of this approach is not to treat criminal offenders as "blameless patients in need of treatment," but instead to recognize that we too share some of the blame, and that society may never impose suffering or deprivation on one of its members except according to the rules of a practice of punishment justified under a public political conception of justice.[35] To impose punishment for any other purpose is to violate the conception of society as a fair system of cooperation between free and equal citizens who are possessed of the moral powers and hold reasonably plural conceptions of the good.

When dealing with epistemic moral failings, the response proposed above recognizes society's role in bringing about the offending behavior and respects our obligation to cultivate one another's moral powers. It also ensures that the basic liberties of all members of society, which are often imperiled by crime, are secured while the offending individual's moral powers are remedially cultivated, or brought up to the level where further criminal epistemic moral failings are unlikely.

Moving on to those cases in which individuals know what justice demands but find themselves unable to comply, the argument proceeds along similar lines. As Jeffrey Howard

---

[35] This means that retribution cannot be offered as a justification unless it can be grounded in a public political conception of justice. Kant's theory of retribution may satisfy this criterion, but that is a topic for future work; the quote is from Howard, Jeffrey W. "Punishment as Moral Fortification." *Law and Philosophy* 36, no. 1 (February 2017): 46.

explains it, those who experience a tension between the demands of justice and their empirical desires have an obligation, rooted in the political conception of justice, to fortify their moral powers.[36] When they experience motivational moral failure, the state may justly intervene to enforce the obligation (fortify their moral powers) pursuant to the political conception of justice.[37]

That individuals have such a duty, Howard argues, is implicit in the way in which we talk about duties. For example:

> To say that Barbara is subject to a duty not to steal from others is not some obtuse shorthand for the idea that someone else needs to ensure that Barbara does not engage in theft […] it means that Barbara needs to ensure that Barbara does not culpably commit theft, […] She must, that is, do what it takes to realize a state of affairs in which she has not culpably committed that crime. To succeed, this may involve, inter alia, avoiding locations in which objects she is tempted to steal are found; reflecting on her potential victims' interests and their value, in the hopes of generating countervailing desires not to wrong them; deepening her understanding of why theft is wrong in the hope that her moral motivation not to steal will be strengthened through such understanding; and so on.[38]

Cultivating and fortifying one's moral powers is part of satisfying the obligations imposed by the political conception of justice. Another example that helps to demonstrate the general obligation one has to prevent oneself from committing crimes is the prohibition against speeding while driving a car. Speeding does not violate anybody's basic liberties, *per se*, but it substantially increases the risk of a violation of these liberties without justification, and thereby violates our duty to take care that we do not commit culpable legal wrongs.[39]

---

[36] Ibid., 45.
[37] Ibid., 46.
[38] Ibid., 47.
[39] Ibid., 48.

We expect society to prevent everyone from committing crimes (in two very particular ways that I will describe below), but we also expect each of us to prevent ourselves from committing crimes by fortifying our moral powers, refraining from negligent behavior, and taking other common measures. This is the basis for the insistence, at least in most common law systems, that all citizens must maintain "reasonable firmness" in the face of criminogenic pressures.[40]

Furthermore, as Howard explains, "it is not a moral oversight that we demand reasonable firmness and yet refuse to permanently curtail citizens' liberty and privacy to ensure that demand is met."[41] Instead, we expect competent individuals to see to their own fortification, and we expect parents and caretakers to see to the fortification of their dependents. The state can appropriately inculcate the competence to fortify one's own first moral power, such as through compulsory primary education, but it would be improper for the state to engage in the sort of continuous monitoring that fortification requires when far less onerous measures can ensure adequate compliance with the law. Furthermore, a system of permanent coercive intervention to promote the first moral power would not afford citizens a fair opportunity to avoid coercive interference in their pursuits of their respective conceptions of the good by complying with the demands of political justice.

What does all of this mean for criminal punishment? Motivational moral failure requires the fortification, rather than the mere remedial cultivation, of the first moral power. The

---

[40] Criminogenic pressures may be emotional (as with crimes of passion) or social (such as norms surrounding underage drinking), and may also include long-term situational factors. For a survey of recent work on criminogenic disadvantage, see Benjamin Ewing, "Recent Work on Punishment and Criminogenic Disadvantage," *Law and Philosophy* 37, no. 1 (February 2018): 29–68.

[41] Howard, "Punishment as Moral Fortification," 51.

offending individual in our hypothetical situation must not be allowed, then, to remain in or enter circumstances where reoffending is likely until the fortification has been successful and they no longer pose an unusual risk to society.[42] An individual with an adequately developed first moral power would submit to these restrictions voluntarily, perhaps even eagerly, but individuals who need fortification most likely will also need to be coercively restrained from reoffending.

Fortification is likely to be more onerous and time consuming than mere remedial cultivation, which has two implications. First, people who are not sufficiently motivated by justice will require a greater incentive to engage in fortification on their own; but the deprivation of liberty imposed by the state to carry out this rehabilitative fortification will also be more onerous, and will likely provide this incentive. Generally speaking, however onerous fortification may be, it is more onerous to undergo it while being deprived of liberty than to undergo it while physically free— which adds an element of incidental deterrence to rehabilitative punishment.[43]

In the final hypothetical, consider somebody who is unabashedly unmoved by concerns of justice, and who commits crimes without regard for the demands of justice.[44] This person does not believe their conduct to be excused or justified, they are simply indifferent to its unlawful character. The amount of rehabilitation such a person would need in order not to pose an unacceptable risk to the basic liberties of others is likely substantial; but the logic under which this person is to be incapacitated and rehabilitated is the same as in the previous hypothetical.

---

[42] In most justified practices of punishment, for reasons analogous to those that I will discuss below with respect to standards of proof, the burden of demonstrating that an individual is not yet ready to be free from coercive fortification will likely rest with society, subject to reasonable presumptions and processes prescribed as part of the practice.

[43] This is important, because it remains an open question whether deterrence alone can justify punishment in a system committed to treating people as ends in themselves.

[44] Often people who fit this description suffer from psychosocial and neurological disorders, but the same logic pertains whether or not this is the case.

There is no room in a political conception of justice for retributive punishment, so the terms and duration of the incapacitation are based on consequential considerations ("when can this person be released without posing an unacceptable risk to society?") and therefore, indirectly, rehabilitative considerations ("how long does it take to ensure that this person will not pose a risk to society when released?"). No person can be subjected to suffering and deprivation unless it can be justified with reference to the public political conception of justice.

On Crime as a Failure of the First Moral Power

So far, the arguments I have advanced for punishment have made two very foundational and related assumptions: the criminal laws that the offenders violated were justified, and the criminal behavior betrayed an insufficiently-developed first moral power. While offering a full account of how substantive criminal prohibitions are justified in political society is beyond the scope of this paper, my reliance on these two assumptions compels me to briefly explore their relationship.

In the previous section, I treated criminal conduct as unjust conduct— as the first moral power applies to appreciating and being moved by the demands of political justice rather than, necessarily, the law. I understand the justified criminal law as those limits on the liberties of individuals necessary to prevent them from interfering with the liberties of others. Rawls' first principle of justice holds that each citizen is entitled to the same basic liberties as set forth in a mutually-compatible scheme. Crimes are, in my conception, those actions that can only be justified in a scheme of liberties that cannot be universalized within a given political society.

Criminal prohibitions, then, serve to actualize the demands of the two principles of justice by placing limits on the liberties of each citizen that render them compatible with all others enjoying the same liberties. Only once a scheme of compatible coextensive liberties exists can a system of offices be established such that there is fair opportunity to gain each office: this includes the office of 'criminal,' to which societies attach considerable disadvantage. Where citizens do not comply with the demands of political justice from respect, they can be made to do so for the benefit of society because they are not behaving as reasonable and rational, and therefore need rehabilitation before they can exercise autonomy.

On the Limits of Permissible Punishment

Before I move on to exploring the justifications for particular impositions of punishment, I want to address one last aspect of the practice of punishment that is often treated on a case by case basis when the problem instead lies at the practice level. As I have repeatedly stated, every societal imposition of suffering or deprivation must be justified by reference to the public political conception of justice. This means that a practice of punishment that prescribes conditions under which suffering and deprivation can be imposed must account for all of the suffering and deprivation that attend the punishments handed down. This not only includes harms to families, communities, and the social fabric, but also includes incidental harms to those condemned, such as sexual violence in prison or the stigma of being branded a criminal.

A just practice of punishment cannot disregard these harms— they must either be justified (if any justification obtains) or eliminated. Where unjustified harms generally or frequently attend certain punishments, society cannot continue imposing those punishments until

the harms have been eliminated. All members of society lend their collective might to the imposition of punishment, so all members of society are responsible for its unjustifiable excesses. No member of society loses their status as a free and equal citizen upon committing, or even being convicted of, a crime— they merely enter a process of rehabilitation, during which they must be treated with dignity, even while they are made to conform to the law.[45]

[45] Obviously, an incarcerated individual is deprived of physical freedom; but the moral freedom relevant to the Rawlsian conception of the citizen transcends physical freedom and is not lessened by criminal offending or conviction. It is the case, however, that an individual with an inadequately-developed first moral power is less morally free than one with a fully-developed first moral power— the former individual is subject to contingent, empirical whims and passions, while the latter is subject only to the self-imposed constraints of justice and is thereby autonomous. This is true irrespective of whether the inadequacy of the first moral power led to any criminal behavior and is independent of any punishment imposed for such behavior upon conviction.

Chapter 4

On Justifying Particular Impositions of Punishment

I have already established that particular impositions of punishment must be justified by reference to the rules that define the justified practice of punishment. Consistent with the demands of political justice, then, I cannot refer directly to moral principles or particular comprehensive conceptions of justice (such as those that call for either retribution or mercy). Instead, I will discuss the sorts of rules that can both be part of the justified Rawlsian practice of punishment and justify particular impositions of punishment.

On the Importance of Legal Guilt and Standards of Proof

The first set of rules that are necessary in justifying particular impositions of punishment are those that pertain to determining legal guilt. Guilt is a prerequisite to punishment in any justified practice of punishment, and so a finding of guilt is a necessary condition for proceeding to subsequent stages of the punishment inquiry. It is important to point out here that I use the term guilt to refer to a legal, rather than natural, fact. While some theories of punishment, including many retributive theories, see punishment as justified by an individual's having engaged in unlawful conduct, our practice of punishment cannot be concerned with what conduct actually took place: instead, we are concerned only with conduct that a defendant can be found guilty of according to the rules of the practice of criminal law, including the practice of punishment.[46]

[46] Were this not so, people could be justifiably punished for engaging in wrongful conduct for which they could not be properly convicted (for lack of admissible evidence, expired

This implies that where individuals are punished for crimes they did not commit, but for which they were duly convicted, the punishment cannot be said to be directly unjust.[47] Instead, wrongful convictions betray flaws and injustices at the level of the practice. Surely there should be a procedure, incorporated into the practice, for challenging and overturning wrongful convictions: representatives in the original position would insist on this for the reasons I return to momentarily. But it would be a mistake to understand wrongful convictions as individual failings, rather than inadequate safeguards to prevent convictions on the basis of inadequate or unreliable evidence, or misconduct by agents of the justice system. Aside from overturning wrongful convictions (because a just practice of law requires it), we should all see wrongful convictions as evidence that we need to reexamine the rules that define our practice and ensure that they are consistent with the demands of justice.

Standards for determining legal guilt vary across legal systems, but a system based on Justice as Fairness would have stringent standards for conviction. People convicted of crimes, especially serious crimes, are subjected to significant suffering and/or deprivation at the hands of the state and, from behind a veil of ignorance, reasonable citizens would insist on a fair opportunity to avoid this hard treatment by conforming their conduct to the law.[48] Taking for granted that no system for assessing factual guilt and innocence can be perfectly accurate,

---

statute of limitations, etc.). While the conclusion that innocent people can be justifiably punished when they have been properly convicted is startling, it is also a necessary consequence on the formal nature of justifying impositions of punishment within a practice of positive criminal law. This is part of why citizens would insist on stringent standards for conviction when establishing the practice.

[47] The same may not be true of the underlying conviction, but that is the subject of a future paper.

[48] This may also generate a legitimate reliance interest, though appealing to this interest is unnecessary here.

representatives establishing rules for determining guilt from the Rawlsian original position would likely choose to err on the side of mistakenly acquitting the guilty rather than mistakenly convicting the innocent. In any event, that is the balance most liberal democracies have struck.

On Determining the Appropriateness of Particular Punishments

Aside from the rules for determining guilt and innocence, there would also be rules for determining what punishments can justifiably be imposed in particular situations. The appropriate type and amount of punishment usually varies based on (1) the offense, (2) the circumstances in which the offense occurred, (3) the victim(s) of the offense, and (4) the characteristics of the offender. The main challenge in prescribing punishments is balancing two perceived normative imperatives: that like cases should be treated alike and that punishments (especially rehabilitative, restorative, and incapacitative punishments) should be tied to the needs of particular offenders and communities.

The apparent contradiction between these imperatives is merely an artifact of some ambiguity in terms. The maxim that like cases ought to be treated alike generally refers to cases that are alike in all relevant respects, and should be interpreted to mean that case outcomes should be based on factors that are relevant according to the rules that define the practice under which they are decided upon. These rules, if they conform to the principles of justice, will not allow punishments to differ based on race or social class, as these factors will be considered irrelevant to the question of what punishment should be imposed.

On the other hand, a just practice will consider relevant those factors that determine which methods of rehabilitation, incapacitation, and/or restoration are most likely to protect

society and promote the autonomy of the offender. In any two cases where all of those factors are the same, the punishment imposed should also be the same; but where those factors are different, there is nothing unjust about tailoring punishments to the circumstances in which they are imposed.

Chapter 5

Conclusion

Contemporary philosophy of punishment has, for the most part, evolved parallel to political philosophy: moving in the same direction but not intersecting. As a result, philosophies of punishment rarely make reference to the political principles that are considered legitimate justifications for other social institutions. In this paper, I have attempted to situate a philosophy of criminal punishment within the framework of a broader political philosophy.

I chose to use Rawls' Justice as Fairness because it is among the most influential and widely-respected theories in contemporary political philosophy. I welcome similar projects, however, that seek to determine what a just system of criminal punishment would look like under other political philosophies. These efforts can help to highlight the strengths and weaknesses of both our philosophies of criminal punishment and our political philosophies.

One essential feature of justifying punishment in the Rawlsian framework is that justifications can only refer to public political conceptions of justice. This prevents any particular comprehensive conception of justice being used to punish someone who does not adhere to it, and respects the two moral powers of every citizen. The public political conception of justice is what anybody would chose for a society, from a menu of all known conceptions of justice, if they (1) conceived of society as a fair system of cooperation between free and equal citizens who are reasonable and rational, and (2) made their selection from an "original position" in which they knew general facts about humans and human society but were unaware of their own conceptions of the good.

In *Justice as Fairness*, Rawls argues that the public political conception of justice consists of the two principles of justice, which essentially (1) guarantee everybody the same

scheme of basic liberties and (2) require that deviations in outcomes result from fair competitions for offices that are advantaged because their primary benefit accrues to the least advantaged members of society. These rules apply to all of the institutions that comprise the basic structure of political society.

The practice of punishment is part of the basic structure of society and is therefore an appropriate direct object of Rawls' two principles of justice. As I show above, society has an obligation to cultivate and fortify the first moral power of each of its members. This means that punishment must both reduce the risk that an offending individual with commit future crimes (by temporary incapacitation, such as through incarceration) and promote that individual's ability to comply with the demands of political justice through rehabilitation. Furthermore, all suffering or deprivation imposed pursuant to a practice of punishment must be justified or eliminated, including suffering or deprivation imposed on people other than the convicted or by people who do not represent the state (such as other prisoners).

Once a justified practice of punishment is in place, particular impositions of punishment can only be justified by reference to the rules that define the practice. The essential precondition for imposing punishment is a finding of legal guilt, and the procedures for determining guilt, as set out in the practice of punishment, should be crafted to avoid wrongful conviction. Otherwise, no citizen would have a fair opportunity to avoid punishment by conforming their conduct to the law.

A practice of punishment also sets out rules stating which factors will be considered relevant to punishment decisions. A just practice of punishment will consider the types of punishments that are likely to rehabilitate the offender, and protect society, in deciding which

punishments to impose. Where all relevant factors are the same, a just practice of punishment will also require that the same punishments be imposed across cases and offenders.

To understand the implications of justifying punishment in this way, it is helpful to leave the Rawlsian ideal society and consider the contemporary United States. Most Americans understand punishment in retributivist terms, and our penal instructions are designed to deliver retributive, rather than rehabilitative or consequentialist, punishments. This likely contributes to the United States having both the highest rate of incarceration in the world and a staggeringly high rate of recidivism. The U.S. Department of Justice itself concedes that "increasing the severity of punishment does little to deter crime," and yet its attorneys regularly seek years-long sentences for relatively minor infractions.[49]

If the American system were to be brought in line with the Rawlsian principles of justice, the aim of punishment would be to rehabilitate offenders and thereby make communities safer. A Department of Justice that conceded that "prisons themselves may be schools for learning to commit crimes," and "prison can exacerbate, not reduce, recidivism" would not seek to continue sending people to prison for increasingly long periods of time.[50] It would expand, rather than continue to cut, educational and job-training programs for inmates that allow them to pursue their own conceptions of the good when they leave prison, and it would make confronting and coming to terms with one's past mistakes an important part of an expanded counseling program.

Most importantly, however, a more-just America would send fewer people to prison and reduce the sentences of those who it does incarcerate. Most time spent in prison today cannot be

---

[49] "Five Things About Deterrence" (National Institute of Justice), accessed April 2, 2019, https://nij.gov/five-things/pages/deterrence.aspx.

[50] Ibid.

justified by a public political conception of justice, and is therefore impermissible. The same is true of the suffering that inmates experience in prison due to sexual violence and excessive disciplinary tactics (such as solitary confinement), as well as of the suffering that inmates' loved ones and communities experience as a consequence of incarceration. To the extent that a society could not incarcerate people without imposing unjustifiable suffering, it could not permissibly incarcerate them at all.

This is a fairly radical way of thinking about criminal punishment, but it is grounded in a well-established and respected theory of political justice that is far from radical. If anything, that the implications of the theory seem radical only demonstrates the magnitude of the divergence between political philosophy and the philosophy of law. My hope is that closing the gap between these traditions will provide criminal-justice reformers with arguments for a more robust and equitable understanding of the purpose of punishment that is grounded in the fundamental tenets of mainstream political liberalism.

BIBLIOGRAPHY

Ewing, Benjamin. "Recent Work on Punishment and Criminogenic Disadvantage." *Law and Philosophy* 37, no. 1 (February 2018): 29–68.

"Five Things About Deterrence." National Institute of Justice. Accessed April 2, 2019. https://nij.gov/five-things/pages/deterrence.aspx.

Flanders, Chad. "Criminals behind the Veil: Political Philosophy and Punishment." *Brigham Young University Journal of Public Law* 31, no. 1 (2016): 83–109.

Hart, H. L. A. *The Concept of Law*. 3. ed. Clarendon Law Series. Oxford: Oxford Univ. Press, 2012.

Howard, Jeffrey W. "Punishment as Moral Fortification." *Law and Philosophy* 36, no. 1 (February 2017): 45–75.

Official Playing Rules Committee. *Official Baseball Rules: 2018 Edition*. Office of the Commissioner of Baseball, 2018.

Rawls, John. *A Theory of Justice: Revised Edition*. Cambridge, MA: The Belknap Press of Harvard University Press, 1999.

——— *Justice as Fairness: A Restatement*. Edited by Erin Kelly. Cambridge, MA: The Belknap Press of Harvard University Press, 2001.

——— "Two Concepts of Rules." *The Philosophical Review* 64, no. 1 (January 1955): 3.

ACADEMIC VITA

EDUCATION

The Pennsylvania State University – Schreyer Honors College – University Park, PA (August 2015 – May 2019)

Master of Arts (MA)

- Major in American Politics, with a focus on state trial courts and judicial behavior
- Minor in Methodology, with coursework and training in statistics, survey design, and statistical programming
- Additional graduate coursework in Philosophy and Bioethics

Bachelor of Arts (BA)

- Majors in Political Science and the Philosophy of Justice, Law, and Values
- Coursework spanning twelve academic departments in four different colleges

Oxford University – Law Faculty – Oxford, England (April 2018 – June 2018)

- Studied terrorist groups' strategic uses of the legal form under the supervision of Prof. Fernanda Pirie
- Affiliated with Centre for Socio-Legal Studies as a "Recognized Student"

SELECTED PROFESSIONAL EXPERIENCE

Centre County Office of the Public Defender
Intern Investigator – Centre County, PA (August 2017 – December 2018)

- Conducted investigations and assisted in case preparation to support the defense of indigent clients, both juveniles and adults, facing serious felony charges
- Interpreted and translated for Spanish-speaking clients

Public Defender Service for the District of Columbia (PDS)
Intern Investigator – Washington, D.C. (May 2017 – July 2017)

- Assisted in investigations by finding and interviewing witnesses, photographing crime scenes, recovering physical evidence, obtaining records and surveillance footage, and writing and serving subpoenas
- Defended clients charged with offenses ranging from misdemeanor assault to first-degree murder
- Assisted in case preparation by drafting motions, preparing exhibits, summarizing scientific and technical information, enhancing case-relevant audio for forensic purposes, and reviewing discovery documents
- Participated in every stage of a criminal case in D.C.

Penn State School of Music
Staff Audio Engineer – University Park, PA (May 2016 – May 2017)

- Recorded Penn State School of Music students, faculty, and ensembles in professional recording studio
- Conducted trainings for students interested in learning how to record music

International Model United Nations Association
Substantive Assistant Director at NHSMUN Conference – New York, NY (August 2015 – March 2016)

- Helped stage the largest Model UN conference in the world as part of a diverse, international team of students
- Managed a three-day UN committee simulation with 130 high-school-student participants

SELECTED AWARDS AND HONORS

- 2019 College of the Liberal Arts Student Marshall
- 2018 Rock Ethics Institute Stand Up Award (for ethical leadership and commitment to service)
- 2018 Penn State University Nominee for both Rhodes and Marshall Scholarships
- Rock Ethics Institute Honors Thesis Fellowship (competitive grant that supports promising research in ethics)
- Phi Beta Kappa

ACADEMIC SERVICE

- Sole undergraduate representative on the search committee for the Dean of the College of the Liberal Arts
- Member of the Paterno Fellows Student Advisory Board
- Member of selection committee for 2019 Rock Ethics Institute's Stand Up Award