A Restorative Theory of Criminal Justice

By

Ryan Edward McSheffrey

A thesis submitted to the Department of Philosophy
In conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
October, 2016

Copyright © Ryan Edward McSheffrey, 2016
Abstract

In this project, I defend a restorative theory of criminal justice. I argue that the response to criminal wrongdoing in a just society should take the form of an attempt to heal the damage done to the community resulting from crime. I argue that the moral responsibilities of wrongdoers as wrongdoers ought to provide the framework for how a just society should respond to crime. Following the work of R.A. Duff, I argue that wrongdoers incur second-order duties of moral recognition. Wrongdoers owe it to others to recognize their wrongdoing for what it is, i.e. wrongdoing, and to shoulder certain burdens in order to express their repentant recognition to others via a meaningful apology. In short, wrongdoers owe it to their victims and others in the community to make amends. What I will deny, however, is the now familiar claim in the restorative justice literature that restoring the normative relationships in the community damaged by criminal forms of wrongdoing requires retributive punishment. In my view, how we choose to express the judgement that wrongdoers are blameworthy should flow from an all things considered judgment that is neither reducible to the judgement that the wrongdoer is culpably responsible for wronging others, nor the judgement that the wrongdoer in some basic sense “deserves to suffer” (or “deserves punishment,” etc.).
Acknowledgements

I wish to express my gratitude to the numerous people who directly supported this project. I especially want to thank my supervisor, Christine Sypnowich, for her unwavering support and encouragement. Without her wisdom and guidance, this project would not have been possible. Many thanks also to the other members of my doctoral committee, Rahul Kumar and Alistair Macleod, for their consistently incisive comments and feedback. Without our many meetings and interactions, this project would have developed quite differently. I also want to thank the other members of my examining committee, Antony Duff and Colin Farrelly, for their challenging and thought-provoking comments and questions.

I am also indebted to those in attendance at the 2015 Queen’s Philosophy Graduate Student Colloquium, the 2016 Queen's Philosophy Department Colloquium, and the Queen’s Philosophy Justice League Thesis Research Group. The dissertation was immeasurably improved by the challenging discussions at those events.

The personal support of my family and friends was crucial to the success of this project. I especially want to thank my parents, Kathy and Jim, for their unconditional love and support, without which I would have lacked the personal strength necessary to endure the many challenges involved along the journey.
# Table of Contents

Abstract ....................................................................................................................................... ii  
Acknowledgements ...................................................................................................................... iii  
Table of Contents ......................................................................................................................... iv  

Chapter 1 – Introduction .............................................................................................................. 1  

Chapter 2 – Consequentialism and punishment ........................................................................ 6  
2.1 – Introduction ....................................................................................................................... 6  
2.2 – Problems with strict consequentialism ........................................................................... 7  
2.2.1 – Problems with utilitarianism ....................................................................................... 10  
2.2.2 – Problems with traditional rehabilitation theory ......................................................... 13  
2.3 – Problems with sophisticated strict consequentialism .................................................... 16  
2.4 – Side-constrained consequentialism ................................................................................ 21  
2.5 – Some further remarks concerning consequentialism and punishment ...................... 25  
2.6 – Conclusion ....................................................................................................................... 27  

Chapter 3 – Desert theory .......................................................................................................... 28  
3.1 – Introduction ....................................................................................................................... 28  
3.2 – What is retributivism? ....................................................................................................... 30  
3.3 – Retributivist themes ....................................................................................................... 35  

3.3.1 – Circular retributivism ............................................................................................... 36  
3.3.2 – Hedonic-suffering retributivism ................................................................................. 37  
3.3.3 – Forfeited-rights retributivism ................................................................................... 39  
3.3.4 – Punitive emotions ..................................................................................................... 42  
3.3.5 – Expressive retributivism ........................................................................................... 44  
3.4 – Retributivist compromise theories explored .................................................................. 47  

3.4.1 – Punishment as public censure plus prudential supplements ........................................ 49  
3.4.2 – Communicative punishment, or “teleological” retributivism ........................................ 56  
3.5 – Some further remarks concerning retributivism ............................................................ 61  
3.6 – Conclusion ....................................................................................................................... 65  

Chapter 4 – Moral responsibility and the restorative justice alternative .................................. 67  
4.1 – Introduction ....................................................................................................................... 67  
4.2 – What is restorative justice? ............................................................................................ 69  
4.3 – The informal model of restorative justice explored ....................................................... 71  
4.4 – A retributivist response to the informal model of restorative justice ........................... 81  
4.5 – Moral responsibility, blame, and restorative justice ...................................................... 89  
4.6 – Conclusion ....................................................................................................................... 108  

Chapter 5 – A restorative justice theory of punishment .......................................................... 111  
5.1 – Introduction ....................................................................................................................... 111  
5.2 – Restorative justice and the duties of offenders .............................................................. 113  
5.3 – Moral recognition and the duties of offenders .............................................................. 115  
5.4 – Duties of moral recognition and the limits of enforcement ........................................... 124  
5.5 – Limits of the restorative justice account ...................................................................... 130  
5.6 – Restorative justice and retributivist themes .................................................................. 134  

5.6.1 – Restorative justice and forfeited rights ..................................................................... 136  
5.6.2 – Restorative justice and punitive emotions ................................................................. 137  
5.6.3 – Restorative justice and deserved censure ................................................................. 139  
5.7 – Conclusion ....................................................................................................................... 141
### Chapter 6 – Defending punishment for the sake of moral recognition

- **6.1 – Introduction**
  - Page 144
- **6.2 – Punishing the already repentant and the predictably defiant**
  - **6.2.1 – Punishing the already repentant**
    - Page 147
  - **6.2.2 – Punishing the predictably defiant**
    - Page 151
  - **6.2.3 – Punishment and the problem of psychopathic offenders**
    - Page 153
- **6.3 – Proportionality, moral reform, and the duties of offenders**
  - Page 158
- **6.4 – Restorative justice and societal protection**
  - Page 164
- **6.5 – Moral recognition and the legitimate authority of the state**
  - **6.5.1 – The feasibility complaint**
    - Page 176
  - **6.5.2 – The proportionality complaint**
    - Page 177
  - **6.5.3 – Moral recognition and the liberal state**
    - Page 178
- **6.6 – Conclusion**
  - Page 181

**Chapter 7 – Conclusion**

- Page 186

**Bibliography**

- Page 189
Chapter 1 – Introduction

In this project, I defend a restorative theory of criminal justice. I argue that the response to criminal wrongdoing in a just society should take the form of an attempt to heal the damage done to the community resulting from crime. The title of the project is a restorative theory of criminal justice, rather than a restorative theory of criminal punishment, since the question of the permissibility of criminal punishment is only one question among others concerning the problem of how a just society ought to respond to criminal wrongdoing. According to a restorative justice approach, in my view, responding to criminal wrongdoing by punishing the wrongdoer is not always the best answer.

Traditionally, the philosophy of criminal justice focuses almost exclusively on justifying the institution of criminal punishment. Consequentialists focus on the benefits of punishing wrongdoers, while retributivists focus on why hard treatment is something that wrongdoers deserve to suffer. However, traditional approaches tend to assume as their mutual point of departure that criminal punishment is always, or at least almost always, the appropriate societal response to crime. The central purpose of this project is to interrogate this supposition. I argue that alternatives to criminal punishment might be appropriate in certain contexts given all relevant considerations. Traditional approaches to the problem of criminal wrongdoing focus almost exclusively on what we ought to do to wrongdoers, rather than on what wrongdoers, by virtue of wronging others, ought to themselves do in response to their own crimes. As I will argue, the moral responsibilities of wrongdoers as wrongdoers ought to provide the framework for
how a just society should respond to crime. It is true that, as a community, we
have a collective obligation to hold wrongdoers in appropriate ways accountable,
and to censure criminal forms of misconduct. However, how we choose to
express blame and how we choose to intervene in the lives of wrongdoers should
be guided, at least primarily, by what we take wrongdoers to owe to others by
virtue of being wrongdoers. Following the work of R.A. Duff, I argue that
wrongdoers incur second-order duties of moral recognition. Wrongdoers owe it to
others to recognize their wrongdoing for what it is, i.e. wrongdoing, and to
shoulder certain burdens in order to express their repentant recognition to others
via a meaningful apology. In short, wrongdoers owe it to their victims and others
in the community to make amends.

What I will deny, however, is the now familiar claim in the restorative
justice literature that restoring the normative relationships in the community
damaged by criminal forms of wrongdoing requires retributive punishment. I
agree with H.L.A. Hart, T.M. Scanlon and others, that retributivism about
punishment constitutes “…a mysterious piece of moral alchemy, in which the two
evils of moral wickedness and suffering are transmuted into good.”¹ I argue that
by appealing to a non-retributivist theory of moral responsibility and blame, we
can avoid the controversial view that the punishment of wrongdoers constitutes
an intrinsic good. Instead, in my view, how we choose to express the judgement

¹ Hart, H.L.A., “Postscript: Responsibility and Retribution,” in Punishment and
Oxford University Press 2011) p.163.
that wrongdoers are blameworthy, and hence ought to be held accountable to others, should flow from an all things considered judgment that is neither reducible to the judgement that the wrongdoer is culpably responsible for wronging others, nor the judgement that the wrongdoer in some basic sense “deserves to suffer” (or “deserves punishment,” etc.). In determining the appropriate response to crime, we should consider not only the culpability of the wrongdoer, but also (among other things) what the wrongdoer owes to others by virtue of her wrongdoing, our standing as a community in relation to the wrongdoer, whether or not the wrongdoer is already remorseful and apologetic, and the likely forward-looking restorative effects of our chosen mode of expression on the wrongdoer, her victim(s), and the community more generally. The restorative justice account I defend is therefore highly ambitious. It sets the bar high for what justice requires of the just society when responding to crime. However, as I hope will become apparent in what follows, this should come as no surprise; the ethics of criminal justice is profound and demanding.

In Chapters 2 and 3 I review the traditional approaches. In Chapter 2, I argue that strict consequentialism, while commendable in its focus on the tangible forward-looking benefits associated with criminal justice institutions, fails to treat wrongdoers and the general public with appropriate respect. Criminal justice is about more than simply good outcomes. It is also about appropriate modes of treatment that respect wrongdoers and others as rational and responsible moral agents who possess basic rights. In Chapter 3, I argue that traditional versions of retributivism, while commendable in their willingness to
address the problems that afflict consequentialist accounts, are confronted by the problem of why wrongdoers deserve hard treatment. Without a compelling account of why wrongdoers deserve hard treatment, retributivism looks problematically barbaric. It is often left unclear why wrongdoers deserve to suffer, why the deserved suffering of wrongdoers ought to be considered an intrinsic good, and why it should be considered the appropriate role for the state to mete out the suffering that wrongdoers supposedly deserve.

In Chapter 4, I introduce the restorative justice alternative to the traditional approaches. I argue that the two dominant views of restorative justice should be rejected. Criminal punishment is not necessarily at odds with restorative justice, nor does restorative justice necessarily require criminal punishment. Instead, I argue that by appealing to a non-retributivist account of moral responsibility and blame, we should pursue a middle ground position whereby holding wrongdoers accountable for the commission of “public wrongs” is the appropriate response to crime on the part of the community. However, the appropriate mode of holding wrongdoers accountable and expressing blame ought to factor in considerations independent of the culpability of the wrongdoer. In Chapter 5, I defend a restorative justice theory of punishment. I argue that the secondary duties of wrongdoers ground the permissibility of imposing hard treatment upon them for the sake of pursuing the ends of restorative justice. It is permissible, although in a highly restricted sense, to coercively intervene in the lives of wrongdoers for the sake of persuading them to recognize the wrongfulness of their criminal acts, and enabling them to express a meaningful apology to their victims and others as
co-members of a shared, normative community. In Chapter 6, I conclude by defending the restorative justice account against some prominent objections in the literature.
Chapter 2 – Consequentialism and punishment

2.1 – Introduction

Punishment necessarily involves the intentional infliction of harm upon human beings. For this reason, the punishment of wrongdoers is a highly controversial topic. As such, the efforts of moral, political and legal theorists to justify and defend the time-honoured practice of punishing criminal wrongdoers may be described as "the problem of punishment."²

In search of a justification for criminal punishment, consequentialists focus on the forward-looking aims or goals associated with criminal justice institutions. In this Chapter, I argue that while consequentialist accounts have certain attractive features, such accounts fail to capture the importance of addressing criminal wrongdoing for what it is, i.e. wrongdoing. In this way, such accounts fail to treat wrongdoers with the respect they are owed as rational and responsible moral agents capable of responding to moral reasons. Furthermore, strict consequentialism about punishment fails to address what I refer to as the problem of the basic rights of wrongdoers: on the one hand, wrongdoers possess basic rights against being manipulated or used “as a means” in certain ways. On the other hand, criminal justice institutions involve important forward-looking benefits, such as the prevention of crime, the restoration or rehabilitation of wrongdoers, fostering reparations between wrongdoers and their communities,

etc. Therefore, in my view, in order to be plausible, a theory of punishment must attempt some form of reconciliation between the basic rights of wrongdoers, i.e. against being manipulated or used “as a means” in certain ways, and the important, perhaps necessary, tangible benefits of a coercively imposed system of criminal law.³

I begin by providing a brief survey of some prominent examples of strict consequentialism about punishment (sections 2-3). I then discuss one prominent version of side-constrained consequentialism (section 4), and conclude with some general remarks concerning the viability of consequentialist approaches to the problem of punishment (section 5).

2.2 – Problems with strict consequentialism

First, it is important to highlight the attractive ingredients associated with strictly consequentialist approaches to the problem of punishment. In search of a justification for criminal punishment, the strict consequentialist emphasizes the forward-looking aims or goals that criminal justice institutions ought to pursue. An obvious candidate for such an aim is the prevention of crime.⁴ However, it is important to distinguish between the various aims or goals that might be pursued

³ To reiterate: while not meant to be exhaustive, the two deficiencies associated with strict consequentialism about punishment that I focus on in this Chapter are as follows: (1) that strict consequentialist accounts fail to capture the importance of treating wrongdoers as moral agents guilty of wronging others and capable of responding to moral reasons, and (2) that strict consequentialist accounts fail to capture the importance of respecting the basic rights of offenders against being manipulated or used “as a means” in various ways, for the sake of achieving some greater societal good, e.g. crime prevention.
by different consequentialist accounts for the sake of crime prevention. As Duff argues, deterrence, incapacitation and rehabilitative or reformative treatment all might be pursued for the sake of attempting to achieve the further valuable goal of preventing crime.\(^5\) Therefore, we should not assume that any single approach (such as a deterrent approach) has a monopoly on crime prevention as a valuable aim that ought to be pursued by a system of criminal punishment.

Punishment is extremely costly. But if a system of punishment is designed with concrete goals in mind, and efficient in the pursuit of such goals, then, the consequentialist maintains, the costs associated with administering a system of criminal punishment may be outweighed by its benefits. Indeed, the strictly consequentialist justification for criminal punishment depends upon precisely such a calculus. Since consequentialism about punishment focuses on forward-looking societal aims or goals such as crime prevention, it seems, at least at first glance, straightforward for the consequentialist to attempt to justify the significant moral and material costs associated with the design and maintenance of a system of criminal punishment.

On the other hand, as will be argued in what follows, strictly consequentialist approaches to the problem of punishment are vulnerable to the same sorts of objections associated with strictly consequentialist approaches to moral and political theory more generally, as when, for example, critics argue that the pursuit of maximizing certain valuable societal aims or goals fails to respect

\(^5\) Duff, *Punishment*, pp.4-5.
the status of individuals (including both wrongdoers and members of society more generally) as moral agents who possess basic rights.\(^6\)

First consider the worry regarding rights: arguably the most famous problem with traditional forward-looking theories of punishment is their failure to address, in a plausible way, the problem of the basic rights of wrongdoers. More specifically, the problem with strict versions of consequentialism about punishment is their failure to specify what renders wrongdoers liable to be used as a means toward the further, beneficial ends of criminal justice institutions. Under normal conditions, a non-consequentialist is likely to deem morally unacceptable any form of treatment that uses an individual or some group of individuals as a means to achieving some further, beneficial societal aim or goal. The worry concerning using individuals “as a means” is famous in contemporary moral and political philosophy.\(^7\) The relevant question, then, is as follows: what makes wrongdoing a special case? If criminal punishment is to be justified with reference to its forward-looking benefits, e.g. crime prevention, rehabilitation, the societal reintegration of wrongdoers, etc., then the pursuit of such aims must be reconciled, in some plausible way, with the status of wrongdoers as individuals who possess basic rights.

A separate worry runs alongside the problem of the basic rights of wrongdoers: strictly consequentialist accounts of the justification for criminal

\(^6\) My argument remains neutral on the question of to what extent the appropriate rationale for (at least a certain set of) individual rights requires appeal to consequentialist considerations for the sake of justification.

punishment typically fail to treat wrongdoers with the respect they are owed as rational and responsible moral agents. Justifying punishment as a means of pursuing the aim of crime prevention via general deterrence is a case in point: the aim of pursuing crime prevention via general deterrence relies on providing wrongdoers as well as the general public with prudential reasons not to offend.

In this sense, the threat of punishment is designed to condition the behaviour of citizens, in the same way that a circus trainer, through intimidation and the use and execution of threats, aims to condition the behaviour of a tiger. Thus, insofar as a system of criminal punishment relates to offenders and the general public in the same way that the circus trainer relates to a tiger, such a system fails to treat offenders and the general public with the respect they are owed as rational and responsible moral agents capable of responding to moral reasons.

2.2.1 – Problems with utilitarianism

First consider utilitarianism about punishment. The utilitarian, like all strict consequentialists, holds that criminal punishment is always bad or “evil,” and

8 Another way of describing this problem is in terms of the importance of decision-making or choice to the justification for criminal punishment: wrongdoers, like everyone else, are capable of rational and moral decision-making. Therefore, the criminal law ought to relate to wrongdoers in a way that respects their status as capable of rational and responsible moral decision-making. This point will be further discussed in the final section of this Chapter.

9 The distinction between specific and general deterrence is relevant, here: specific deterrence refers to providing the wrongdoer with prudential reasons not to offend, while general deterrence refers to providing the general public with prudential reasons not to offend, i.e. out of fear of the harm or burden of the criminal sanction.

thus may only be justified with reference to the overriding good consequences it is likely to produce.\textsuperscript{11} For example, Bentham argues,

The general object which all laws have, or ought to have, in common, is to augment the total happiness of the community; and therefore, in the first place, to exclude, as far as may be, every thing that tends to subtract from that happiness: in other words, to exclude mischief. But all punishment is mischief: all punishment in itself is evil. Upon the principle of utility, if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil.\textsuperscript{12}

For Bentham, the “greater evil” that can be avoided by means of punishing wrongdoers is the proliferation of crime or “mischief,” and so crime prevention, on this view, is understood as the principal justifying aim of a system of criminal law and punishment. According to the principle of utility, for Bentham, punishment is justifiable only to the extent that it maximizes utility, understood in terms of overall promotion of pleasure and avoidance of pain, via the prevention of crime.

Arguments designed to discredit Benthamite utilitarianism about punishment are famous in contemporary moral and political philosophy. Since utility-maximization is not sensitive to individual rights (understood as constraints on society’s legitimate pursuit of social goods) the utilitarian view would allow (at least under certain conditions) for the ‘punishment’ of the innocent or the disproportionately harsh punishment of the guilty.\textsuperscript{13} Along similar lines, pursuing crime prevention as a way of maximizing utility potentially provides wrongdoers as well as the general public with what are, strictly speaking, morally irrelevant

\textsuperscript{12} Bentham, "Punishment and Utility," p.68.
reasons not to offend. As such, the traditional utilitarian account is vulnerable to
the objection that it treats people as something less than fully rational and
responsible moral agents, i.e. in the same way that the circus trainer aims to
condition the behaviour of a tiger.\footnote{See von Hirsch, \textit{Censure and Sanctions}, p. 11. Cited above note 10.} Failing to take seriously the moral status of
the individual as a rational and responsible moral agent, therefore, is at the heart
of the most common objection to utilitarianism in general, and utilitarian
punishment in particular.\footnote{Duff, \textit{Punishment}, p.3.}

In my view, all versions of strict consequentialism about punishment – i.e.
any view holding that what justifies punishing wrongdoers is \textit{nothing other than}
the successful promotion of the aims or goals of criminal punishment institutions
– are vulnerable to this sort of objection: in the pursuit of maximizing some form
of socially valuable end, e.g. crime prevention, the strict consequentialist ignores
the important point that criminal punishment institutions ought to address both
the welfare of the community as a whole, so to speak, and the welfare and
individual rights of all the individual members of the community, including
wrongdoers. As Rawls famously writes, “Each person possesses an inviolability
founded on justice that even the welfare of society as a whole cannot override.”\footnote{Rawls, \textit{A Theory of Justice}, p.3.}

Presumably, this claim also applies in some way to wrongdoers. Therefore, the
structure of any strictly consequentialist theory of punishment is vulnerable to the
objection that such an account would, under certain conditions, permit us to set
aside or ignore the fundamental inviolability of individuals that Rawls describes,
for the sake of whatever societal goods are specified by the theory in question. The upshot is that the central challenge for utilitarian and other strictly consequentialist accounts that focus exclusively on the prevention of crime is to attempt to reconcile, in some plausible way, the pursuit of crime prevention with the moral status of wrongdoers as rational and responsible agents who possess basic rights.

2.2.2 – Problems with traditional rehabilitation theory

The traditional rehabilitative account is often construed as a strictly consequentialist approach to the problem of punishment.\(^{17}\) On this view, restoration, rehabilitation and the societal reintegration of wrongdoers are understood to be the central justificatory aims criminal justice institutions ought to pursue. However, in contrast with utilitarianism and other forms of strict consequentialism about punishment, traditional rehabilitation theory holds that wrongdoers are in some sense morally ill, and hence, as opposed to being subjected to punishment per se, ought to become the subjects of reformatory or rehabilitative treatment programs that make use of the techniques of modern behavioural and social science.\(^{18}\) For example, Menninger writes,


If we were to follow scientific methods, the convicted offender would be contained indefinitely pending a decision as to whether and how and when to reintroduce him successfully into society. And the skill and knowledge of modern behavioural science would be used to examine his personality assets, his liabilities and potentialities, the environment from which he came [etc.]... Having arrived at some diagnostic grasp of the offender’s personality, those in charge can decide whether there is a chance that he can be redirected into a mutually satisfactory adaptation of the world.\(^\text{19}\)

On this view, therefore, the wrongdoer is understood as a moral patient, and wrongdoing is understood as the manifestation of a maladaptation on the part of the wrongdoer to the moral and social environment.

Like the traditional utilitarian theory of punishment, the traditional rehabilitative view has fallen from grace in the contemporary philosophical literature. Critics cite both the practical and moral failings of any comprehensively therapeutic or corrective approach to the problem of criminal wrongdoing. In practice, rehabilitative treatment often fails to achieve its desired results.\(^\text{20}\) More importantly for my purposes is the critique based on moral principle. Critics have denounced traditional rehabilitation theorists for adopting a view that involves the medicalization of wrongdoing, arguing that it is in an important sense disrespectful to treat wrongdoers as necessarily ill, as it robs wrongdoers of their status as rational and responsible moral agents.\(^\text{21}\) The Achilles heel of the traditional rehabilitative account, therefore, is that it treats wrongdoers as something less than fully rational and responsible. Indeed, such accounts often describe the moral status of wrongdoers explicitly in terms of patienthood, and


\(^{21}\) Murphy, “Marxism and Retribution,” pp.28-9.
coercive treatment programs in terms of molding patients into compliance with societal norms. But, critics argue, wrongdoers, like everyone else, ought to be treated only in ways that respect their status as fully rational and responsible moral agents. Any purely “patient-centered” approach to criminal justice, therefore, systematically fails to treat wrongdoers with the respect they are owed, and therefore ought to be denounced as morally unacceptable.\textsuperscript{22}

Arguments against the traditional rehabilitative view are numerous and well known in the literature. Most of the well known objections to traditional corrective approaches make reference to the tension between coercive, corrective treatment and the moral status of wrongdoers as rational and responsible moral agents. As Murphy writes, “The therapeutic state, where prisons are called hospitals and jailers are called psychiatrists … raises … problems about the justification of coercion and its reconciliation with autonomy…”\textsuperscript{23} Similarly, Moore argues,

\begin{quote}
[\textit{R}ecasting … punishment in terms of ‘treatment’ for the good of the criminal makes possible a kind of moral blindness that is dangerous in itself. … [A]dopting such a ‘humanitarian’ conceptualization of punishment makes it easy to inflict treatments and sentences that need bear no relation to the desert of the offender.\textsuperscript{24}
\end{quote}

Such criticisms are damning indeed.

Alongside these worries, defenders of the traditional rehabilitative view also face the objection that such an account fails to address in a plausible way the problem of the basic rights of wrongdoers: even if the arguably laudable aims

\footnotesize
\textsuperscript{22} For this point see Murphy, “Marxism and Retribution,” pp.6-7.
\textsuperscript{23} Murphy, “Marxism and Retribution,” p.29.
of reformation, rehabilitation and societal reintegration replace the aim of crime prevention via general deterrence, such an account must nevertheless explain what renders wrongdoers liable to be subjected to coercive treatment for the sake of pursuing such ends. Therefore, like defenders of utilitarianism about punishment, the challenge for advocates of coercive corrective or reformative modes of criminal treatment is to attempt to reconcile, in some plausible way, such coercive modes of criminal treatment with the moral status of wrongdoers not just as rational and responsible agents, but also as individuals who possess basic rights against being manipulated for the sake of some greater societal good.

Nonetheless, the traditional rehabilitative model is not without merit. It is widely acknowledged in the literature and in many real-world systems of criminal justice that we should find some role, however limited, for reformative modes of criminal treatment within the context of criminal justice institutions. Despite the problems mentioned above, in my view, the restoration and societal reintegration of wrongdoers are important forward-looking aims associated with criminal justice institutions that should not be dismissed out of hand. In later Chapters (particularly Chapters 4, 5 and 6) I return to the problem of how such aims might be reconciled with the status of wrongdoers as rational and responsible moral agents who possess basic rights.

2.3 – Problems with sophisticated strict consequentialism
As previously observed, the challenge for strictly consequentialist views is to attempt to reconcile, in some plausible way, the beneficial aims of criminal punishment with the moral status of wrongdoers as rational and responsible agents who possess basic rights. More recent consequentialist accounts attempt to address this challenge head on. Therefore, it is worth briefly explaining and addressing one such account.

In my view, sophisticated versions of strict consequentialism are commendable for attempting to take seriously both the status of wrongdoers as moral agents, and the problem of the basic rights of wrongdoers. However, as will be argued in what follows, the structural features of sophisticated versions of strict consequentialism render such accounts vulnerable to the objection that they protect the basic rights of wrongdoers in the wrong way, or for the wrong reasons.

Braithwaite and Pettit’s “republican” theory of criminal justice is a case in point. Braithwaite and Pettit situate their theory of criminal justice within the context of a wider approach to political theory. Their account is in some ways compelling, especially with respect to their explicit attempt to avoid the problems associated with the more traditional strictly consequentialist accounts previously

26 Ibid. Unfortunately, for the sake of brevity, I am not able to address in detail Braithwaite and Pettit’s more general views about the appropriate justification for political authority. What I want to focus on in what follows, rather, is their contention that an approach to criminal justice ought to be both strictly consequentialist and yet, within such a framework, include adequate constraints designed to protect the basic rights (understood in terms of “dominion,” as will be explained in what follows) of offenders.
discussed. They argue that the republican value of “dominion”\textsuperscript{27} ought to be considered the single aim or goal that criminal justice institutions ought to seek to maximize, instead of the familiar consequentialist aim of crime prevention. Nonetheless, in keeping with the spirit of a strictly consequentialist approach, they argue that criminal justice institutions ought to seek “the maximization of the dominion of individual people.”\textsuperscript{28} Dominion, for Braithwaite and Pettit, is preferable to crime prevention as a penal aim, in part since aiming to maximize dominion will not prescribe punishing the innocent, nor prescribe “punishments that exceed uncontroversial limits in degree or kind….”\textsuperscript{29} This is in part since dominion involves taking into account both the direct and indirect implications of penal policy: the dominion of those directly affected by penal policy (e.g. wrongdoers) must be weighed alongside those indirectly affected (e.g. the general public), since maximizing dominion takes into consideration not just the direct dominion-infringing effect of punishment on the wrongdoer, but also the public’s assurance that their freedom (understood as dominion) will not be unduly infringed by whatever criminal justice policies are adopted by the state. Therefore, punishing the innocent, for example, is ruled out, since it will undoubtedly negatively affect the public’s assurance in the requisite sense. The same thing goes for disproportionately harsh forms of criminal punishment. Punishing wrongdoers with undue severity would involve, according to

\textsuperscript{27} Braithwaite and Pettit define dominion as the republican as opposed to the liberal notion of negative liberty. As Duff nicely explains, for Braithwaite and Pettit, dominion is best understood as “…the assured and equal freedom of citizens under the law.” Duff, “Penal Communications,” p.21.

\textsuperscript{28} My italics. See Braithwaite and Pettit, \textit{Not Just Deserts}, p.54.

\textsuperscript{29} See Braithwaite and Pettit, \textit{Not Just Deserts}, pp.78-9
Braithwaite and Pettit, “...inflicting a certain and grievous damage on dominion ... that ... is unlikely to be offset by an appropriately large increase ... in the level of overall dominion...”

However, Braithwaite and Pettit’s account is deficient in an important respect: their account does not provide the right sort of protection for the basic rights of offenders. According to their argument, the aim of dominion maximization is consistent with protecting the basic rights of offenders so that they are not used as a mere means only insofar as certain empirical assumptions are satisfied, e.g. that the public’s assurance will be negatively affected in the requisite sense by certain rights-infringing (or in Braithwaite and Pettit’s terms, dominion-infringing) penal policies. It might turn out that in some particular context, however, the public’s assurance would not be negatively affected in the requisite sense, or not as negatively affected as Braithwaite and Pettit suggest.

It is inappropriate to rely on a constraint designed to protect the basic rights of offenders that is established merely by appeal to a contingent set of facts, or in other words, merely by appeal to how the empirical data are likely to play out in practice. Another way of putting the point is as follows: even if Braithwaite and Pettit are right that scapegoating the innocent or punishing the guilty with disproportionate severity will always have the sort of negative effect on overall levels of dominion that they suppose, it seems to capture the basic rights of offenders in the wrong way, or for the wrong reasons. As Duff eloquently explains,

30 Braithwaite and Pettit, Not Just Deserts, p.79.
While I am not clear how we are to carry out the calculus that weighs the dominion-infringing against the dominion-protecting effects of different penal policies or acts, I do not see how ... [Braithwaite and Pettit’s] confidence that that calculus will result in firm protection for the innocent against being framed, or for the guilty against punishments that we would regard as clearly excessive, is justified: they still face ... the familiar objection to any purely consequentialist theory, that it makes such protections for the individual contingent on the likely effects of particular policies in particular contexts and, thus, vulnerable to infringement when it would be useful to sacrifice the individual for some greater social good.31

Therefore, we should be deeply suspicious of sophisticated versions of strict consequentialism about punishment for some of the same reasons we should be deeply suspicious of more traditional versions of strict consequentialism about punishment: such accounts fail to capture in a plausible way the importance of treating wrongdoers in ways that respect their status as individuals who possess basic rights.

The above arguments suggest that an appropriate justification for criminal punishment must take into consideration more than merely the outputs of criminal justice institutions. If a system of criminal punishment is to be justified, it must also incorporate considerations concerning the moral status of wrongdoers and members of the general public, and the appropriate treatment of wrongdoers and members of the general public with respect to their moral status in the context of criminal justice. In my view, therefore, any plausible forward-looking theory of punishment must include some form of non-consequentialist desiderata concerning what ought to count as permissible criminal treatment in the context of criminal justice.

31 Duff, “Penal Communications,” p.22.
2.4 – Side-constrained consequentialism

As will be discussed in detail in Chapter 3, the traditional rival of strict consequentialism about punishment is retributivism, or “desert theory.”\(^{32}\) Traditionally, those who deny that punishment ought to be justified (at least primarily) with reference its forward-looking benefits, turn their attention instead to backward-looking moral considerations, such as the culpability or desert of offenders. Traditionally, retributivists argue that by virtue of wronging others, wrongdoers deserve to be punished. According to the desert-based view, punishment is something that is valuable intrinsically, or without reference to the forward-looking benefits associated with criminal justice institutions.\(^{33}\) On this view, what justifies punishing wrongdoers is simply that justice demands it. As will be discussed in Chapter 3, traditional retributivist accounts are in this way able to overcome the worries associated with strict consequentialism outlined above. However, retributivists face a distinct set of formidable objections.

With the distinction between consequentialism and retributivism in hand, we are in a position to introduce a third approach: compromise theories attempt to incorporate both forward-looking and backward-looking considerations into their specified justification for punishing wrongdoers. Defenders of “mixed” theories appeal to both the forward-looking benefits of criminal justice institutions, e.g. crime prevention, as well as some form of non-consequentialist desiderata

\(^{32}\) For example see Moore, *Placing Blame*, pp.87-8. Moore writes, “Retributivism … is the view that punishment is justified by the desert of the offender.”

\(^{33}\) *Ibid.*
for permissible punishment. Compromise theories hold that some form of backward-looking criteria for permissible punishment ought to be considered a constraint on the pursuit of whatever beneficial forward-looking aims are specified by the theory. One obvious candidate for such a constraint is the desert of offenders.\(^{34}\)

Importantly, however, not all defenders of a mixed theory approach pick out the desert of offenders as the appropriate backward-looking constraint on the pursuit of the forward-looking aims of a system of criminal punishment. H.L.A. Hart’s famous “rule consequentialist” account is a case in point. Hart argues that “confusing shadow-fighting” between utilitarians and retributivists may be avoided if all parties recognize the consistency of the following two claims: (1) “…that the General Justifying Aim of the practice of punishment is its beneficial consequences...” and (2) “…that the pursuit of this General Aim should be qualified or restricted out of deference to principles of Distribution which require that punishment should be only of an offender for an offence.”\(^{35}\)

For Hart, crime prevention is the principal “beneficial consequence” of a system of criminal law.\(^{36}\) But, according to Hart, we should qualify or restrict our pursuit of this fundamental aim with reference to “principles of Distribution” (e.g., the principle that guilt is a necessary condition for criminal liability), in this way

\(^{34}\) Mixed theories that appeal to the desert of offenders as a constraint on permissible punishment are often referred to as forms of “negative” retributivism. The distinction between so-called negative and positive versions of retributivism will be discussed in more detail in Chapter 3.


constraining our pursuit of crime prevention by means of the criminal law. This avoids the objection, Hart argues, that strictly consequentialist justifications for punishment would, under certain conditions, allow for the punishment of the innocent. But, one might ask, how does Hart account for the liability of offenders to suffer punishment for the sake of crime prevention? While Hart’s view clearly rules out the punishment of the innocent, a further question concerns what renders wrongdoers liable to be used for the sake of pursuing the end of crime prevention. In other words, one might ask, how does Hart’s account address the problem of the basic rights of wrongdoers?

On one possible construal, Hart’s answer lies in the ability of the wrongdoer to predict her future, in the sense of being able to predict how she is likely to be treated by the coercive apparatus of the criminal law, if and when she decides to break it. On this view, therefore, a principle of voluntariness is what renders wrongdoers liable to be used by a system of criminal punishment for the sake of crime prevention. By voluntarily breaking the law, the wrongdoer in a sense consents or implicitly acquiesces to the legal consequences of her action, insofar as such consequences are clear and determinate.

37 Hart, “Prolegomenon to the Principles of Punishment,” p.22.
But this view is clearly unacceptable. It would reduce the normative significance of criminal liability to whatever the law happens to in fact prohibit. In this way, on such a view, if a system of criminal law were in itself abhorrently unjust (say, if it were blatantly racist, or otherwise grossly discriminatory), then the same argument might be put forth as a justification for why wrongdoers ought to be considered liable to suffer deterrent punishment (albeit within the context of an abhorrent system of criminal law). Clearly, we should not accept a theory of punishment that would consider wrongdoers who voluntarily break the law, however grossly discriminatory the law happens to be, morally liable to suffer punishment. Therefore, the principle of voluntariness fails to be convincing.

However, we should not be too hard on Hart’s account, at least with respect to its structural features. In my view, while Hart’s account is substantively incorrect, Hart is correct in pointing out that a fundamentally forward-looking theory of punishment can and ought to be defended, so long as it is constrained by some form of non-consequentialist desiderata for what ought to count as permissible punishment. The problem for Hart’s view, I submit, is merely that he fails to specify a plausible account of the appropriate non-consequentialist desiderata for permissible punishment.

---

39 For this objection to Hart’s view, see Duff, “Penal Communications,” p.13.
40 Also, I will argue that the forward-looking aim of restoration or reparation of the normative relationships damaged by criminal wrongdoing ought to replace the forward-looking aim specified by Hart’s account, i.e. the prevention of crime, although, as I will argue, in my view crime prevention is not irrelevant to the appropriate forward-looking restorative aims of criminal justice institutions. This point will be discussed in Chapter 6.
2.5 – Some further remarks concerning consequentialism and punishment

Thus far, I have focused on two associated problems with consequentialist accounts: (1) a system of criminal punishment must respect the status of wrongdoers as rational and responsible moral agents. Wrongdoers are moral agents, rather than mere moral patients. Therefore, criminal justice institutions must address wrongdoers as rational and responsible agents capable of responding to moral reasons. Also, (2) a justification for criminal punishment must address the problem of the basic rights of wrongdoers. Wrongdoers possess what Rawls describes as a fundamental “inviolability founded on justice that even the welfare of society as a whole cannot override.” In order to be plausible, therefore, a theory of punishment must attempt some form of reconciliation between the basic rights of wrongdoers against being used “as a means” in certain ways, and the tangible, forward-looking benefits of a coercively imposed system of criminal law, e.g. the prevention of crime.

What unites these problems, I submit, is that criminal wrongdoing is not identical with other societal problems that might be addressed by institutional efforts aimed at social engineering, e.g. healthcare, education, poverty relief, promoting general welfare or flourishing, etc. An institution of criminal punishment is a response to a distinct form of societal problem, namely, the problem of criminal wrongdoing. One way of understanding the various deficiencies associated with strict consequentialist accounts is that they tend to view the problem of criminal wrongdoing, and hence the appropriate societal

---

response to the problem of criminal wrongdoing, as identical with other social problems that do not necessarily involve individual citizens culpably wronging others. For this reason, consequentialists have been criticized for eliminating or diminishing the moral importance of the culpability of offenders. If culpable wrongdoing is not a condition for permissible punishment, or permissible degrees of punishment, then the consequentialist both fails to treat wrongdoers with the respect they are owed as rational and responsible moral agents, i.e. as wrongdoers, and also fails to treat wrongdoers and the general public in ways that respect basic rights. Intuitively, both “punishing” the innocent and punishing wrongdoers with disproportionate severity seem to be examples of rights violations, no matter what social goods might be instrumentally related to such forms of treatment. Therefore, a theory of punishment, in order to be plausible, must in some way incorporate the culpability of offenders as a condition for permissible punishment. Furthermore, this suggests that criminal justice is about more than merely good outcomes. It is also about the appropriate treatment of wrongdoers and the general public as a matter of justice.42

Of course, the most famous attempt to incorporate the culpability of offenders as a condition for permissible punishment is to adopt a retributivist theory of punishment. If criminal punishment is something that wrongdoers (in

42 For example see Bennett, C., *The Apology Ritual: A Philosophical Theory of Punishment* (Cambridge: Cambridge University Press, 2008) pp.18-19. These observations might lead one to suppose that retributivism is the only appropriate way to address the problem of criminal wrongdoing. However, in the remainder of this thesis (particularly in Chapters 4, 5 and 6) I will interrogate this claim, and argue that the duties of offenders provide a plausible alternative to retributivism as a way of incorporating the moral importance of the culpability condition.
some sense) deserve to suffer, then permissible punishment ought to track the culpability condition, both in terms of who may be permissibly punished (i.e. only punishing those who deserve it), and in terms of how much wrongdoers may be permissibly punished (i.e. punishing wrongdoers only to the extent that they deserve it).

2.6 – Conclusion

In this Chapter, I argued that consequentialist theories of punishment fail to treat wrongdoers with the respect they are owed both as moral agents and as individuals who possess basic rights. In the next Chapter, I discuss various versions of retributivism. My aim in Chapter 3 will be to unpack several well-known versions of retributivism, both in order to highlight their respective strengths and weaknesses, and to situate in context the view I aim to subsequently articulate and defend.
Chapter 3 – Desert theory

3.1 – Introduction

Traditionally, the philosophy of punishment largely consists of two opposing sides: on one side, consequentialists argue that punishing wrongdoers has important forward-looking benefits, and therefore that punishment, although regrettable insofar as it inflicts pain, ought to be considered permissible for the sake of achieving the ends or goals that an institution of criminal punishment is likely to yield, e.g., crime prevention, social cohesion, the restoration or rehabilitation of wrongdoers, etc.43 On the other side, retributivists argue that punishment is appropriate (or ‘right,’ or ‘just’) independently of its further intended aims or goals, since wrongdoers in some sense deserve to be punished. Retributivists argue that punishment ought to be considered permissible (or indeed obligatory) with reference to its intrinsic value, or its value unrelated to ends beyond the punishment of wrongdoers itself.44 On this view, the practice of criminal punishment is often defended as a way of respecting the status of wrongdoers as rational and responsible moral agents.45 More recently, a third approach has emerged in the literature: compromise theories attempt to bridge the gap between exclusively forward-looking and exclusively backward-looking

43 For the most famous version of this view, see Bentham, J. "Punishment and Utility," p.68. Cited above note 12.
44 As I suggest in what follows, the notion that the punishment of wrongdoers is an intrinsic good is what is essential to retributivism about punishment. For this point see Moore, Placing Blame, p.157.
45 For this point, see Murphy, “Marxism and Retribution,” p.17.
accounts. Compromise theories involve justifying punishment with reference to forward-looking aims, but typically limit or constrain the justifiable pursuit of such aims with reference to some form of non-consequentialist desiderata for permissible punishment.\(^{46}\)

In Chapter 2, I argued that traditional versions of strict consequentialism fail to respect the status of wrongdoers as rational and responsible moral agents capable of responding to moral reasons. Furthermore, such accounts, including more sophisticated versions, fail to adequately address the problem of the basic rights of wrongdoers. The challenge for fundamentally forward-looking theories of punishment, therefore, is to attempt to reconcile, in some plausible way, the basic moral status (i.e. as rational and responsible moral agents) and basic rights (i.e. against being manipulated or used “as a means” in certain ways) of wrongdoers with the pursuit of the various forward-looking aims of criminal justice institutions emphasized by such accounts.

In this Chapter, I describe and review some well-known retributivist theories of punishment. I argue that traditional versions of retributivism are valuable insofar as they address the problems that confront consequentialist accounts (section 2). However, the principle of desert underlying retributivist accounts tends to be elaborated in a way that is either left objectionably vague, or else substantively objectionable (section 3). In short, retributivists fail to explain in a compelling way why the desert of offenders justifies the imposition of

\(^{46}\) Subsequently, in the remainder of this thesis (particularly in Chapters 5 and 6) I defend a non-retributivist, forward-looking compromise theory. For a structurally similar account, i.e. a fundamentally forward-looking compromise theory, see Hart, “Prolegomenon to the Principles of Punishment,” p.9. Cited above note 35.
punishment upon them. Subsequently, I aim to unpack and review two sophisticated retributivist compromise theories (section 4), in order to set the stage for the restorative justice alternative that will be the focus of Chapter 4. I conclude with some general remarks concerning the viability of retributivist approaches to the problem of punishment (section 5). In what follows, rather than attempt to refute the various available versions of retributivism, my aim is to highlight certain strengths and weaknesses associated with such accounts, in order to situate in context an alternative, restorative justice approach that I aim to subsequently articulate and defend in the remainder of the thesis.

3.2 – What is retributivism?

Retributivism is the view that the desert of offenders supplies the justification for criminal punishment. In other words, for retributivists, criminal punishment is not justified with reference to the further goods or benefits that might thereby be produced by punishing wrongdoers. Rather, retributivists argue that the punishment of wrongdoers ought to be considered an intrinsic good, i.e., something that is good in some sense in itself. On this point, Immanuel Kant argues,

Punishment ... can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society. It must always be inflicted upon him only because he has committed a crime...

Furthermore, for Kant,

The law of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaimonism in order to discover
something that releases the criminal from punishment or even reduces its amount by the advantage it promises...\(^{47}\)

Along similar lines, in the contemporary context, the famous retributivist M.S. Moore argues that

Retributivism … is the view that punishment is justified by the desert of the offender. The good that is achieved by punishing, on this view, has nothing to do with future states of affairs, such as the prevention of crime or the maintenance of social cohesion. Rather, the good that punishment achieves is that someone who deserves it gets it. Punishment for the guilty is thus for the retributivist an *intrinsic* good, not the merely *instrumental* good that it may be to the utilitarian or the rehabilitative theorist.\(^{48}\)

Therefore, retributivists are committed to the claim that criminal punishment is justified by the desert of the offender, or in other words, without reference to the further goods to which punishment might be instrumentally related.\(^{49}\)

As such, one attractive aspect of traditional retributivism is that it takes as its point of departure an effort to address wrongdoers as rational and responsible moral agents. If punishment is something that wrongdoers deserve, rather than merely a technique for bringing about certain valuable aims or goals (either for

---


\(^{48}\) See Moore, *Placing Blame*, p.87-88 (Moore’s italics).

\(^{49}\) In my view, while retributivism has certain weaknesses, as mentioned above I will not offer in what follows a comprehensive argument against the view that the punishment of wrongdoers ought to be considered intrinsically valuable (or valuable ‘in itself’), full stop. Instead, it is important for my purposes to outline the basic features of any retributive theory of punishment for the following reasons: in the remainder of the thesis, I shall make use of retributivism as a frame of reference in order to (1) clearly identify the ways in which the view I aim to subsequently defend is *not* a version of retributivism, and in order to (2) highlight the features of the view I aim to defend that are able to capture certain attractive ingredients typically associated with traditional versions of retributivism, i.e. that punishment ought to be “for an offence,” and ought to address wrongdoers as rational and responsible moral agents who possess basic rights.
the wrongdoer or for society), then criminal punishment on such a view aims to treat wrongdoers in a way that respects their status as rational and responsible moral decision-makers.50 On this point, Murphy argues, “…one virtue of the retributive theory … is that it manifests at least a formal or abstract respect for rights, dignity, and autonomy. For it at least recognizes the importance of attempting to construe state coercion in such a way that it is a product of each man’s rational will.”51

Importantly, defenders of retributive punishment typically adopt a ‘deontological’ or ‘rights-based’ approach to ethics and political theory more generally. This is also an attractive characteristic of traditional retributivism since, as argued in Chapter 2, any plausible theory of punishment must account in some way for the problem of the basic rights of wrongdoers. If it can be shown that wrongdoers deserve to suffer punishment, then pursuing the further aims or goals to which a system of punishment might be instrumentally related arguably ceases to be a problem, since the liability of wrongdoers to suffer punishment is established via the principle of desert. The retributivist might argue that if the deserved punishment of wrongdoers is also instrumentally related to further valuable outcomes, then so much the better.52

50 As distinct, we might say, from the case of the circus trainer and her tiger. See von Hirsch, Censure and Sanctions, p.11. Also see Chapter 2, section 2.
51 Murphy, “Marxism and Retribution,” p.17.
52 For this argument, see Nozick, R., Philosophical Explanations (Cambridge, Massachusetts: Harvard University Press, 1981) p.374. Nozick argues that the “further consequences” of criminal punishment beyond the act of taking retribution (understand in Nozick’s terms as connecting the wrongdoer with correct values) “…are not to be dismissed simply…” Rather, he argues, “…we
However, as will be discussed in what follows, traditional retributivism also involves certain troublesome characteristics. Commentators have pointed out that, for retributivists, the desert principle that underlies the liability of wrongdoers to suffer punishment is both theoretically and practically controversial. Many prominent contemporary moral philosophers have argued that, after taking into account a plausible view of freedom and its relationship to moral responsibility, it becomes apparent that no one, including wrongdoers, ought to be considered as deserving of suffering. In a more practical context, some have argued that even if a retributivist construal of the appropriate justification for punishment has moral credibility, a further argument is necessary concerning why it ought to be considered the appropriate role of the state per se to mete out the suffering that wrongdoers supposedly deserve. At the very least, therefore, the desert-based justification for punishing wrongdoers that is essential to retributivism remains highly controversial.

There are, however, a variety of retributivist theories. Positive retributivists argue that wrongdoers deserve to be punished, and hence that criminal

\]

\[\text{\footnotesize 54 For a version of this point, see Duff, “Penal Communications,” p.27. According to one version of this complaint, setting up a costly set of public institutions to bring about the deserved suffering of wrongdoers remains highly controversial in part since, in a world of scarcity such as our own, the aim of meting out the suffering that wrongdoers supposedly deserve must, after all, compete for scarce resources with the pursuit of other valuable societal aims, some of which are directly or indirectly related to the problem of criminal wrongdoing, e.g. healthcare, education, poverty relief, promoting general welfare or flourishing, etc. See Bennett, \textit{The Apology Ritual}, pp.14-15.}\]
punishment is the appropriate backward-looking response to criminal wrongdoing, full stop. In this sense, positive retributivists view retribution (or the act / process of taking retribution) as a positive reason to punish wrongdoers. Negative retributivists, by contrast, argue that the practice or use of criminal punishment (to achieve further aims) ought to be restricted or constrained with reference to the desert of offenders. In this sense, negative retributivists provide reasons why certain individuals should not be punished (e.g. those who do not deserve it, those who deserve less of it, etc.), and thus negative retributivists espouse a view that opposes any purely consequentialist or forward-looking account of the appropriate justification for criminal punishment.55

But, one might ask, to what extent are these positive and negative versions both retributivist accounts of the justification for punishment? In other words, what features do all versions of retributivism about punishment have in common, qua retributivism? The appropriate answer to this question will be grounded in a conceptual debate about how to best understand the concepts “retribution” and “retributivism.” In my view, the essential claim all retributivists seem committed to, by virtue of being retributivist, is Moore’s claim that criminal punishment is, in some sense, intrinsically valuable, i.e. valuable in itself, or without reference to ends beyond the punishment of the wrongdoer itself.56 As Moore correctly argues, “...what is distinctively retributivist is the view that the

55 For a discussion of the distinction between positive and negative retributivism, see Duff, Punishment, p.19.
56 See Moore, Placing Blame, p.157.
guilty receiving their just deserts is an intrinsic good.”\footnote{Ibid (Moore’s italics).} Therefore, if a version of so-called negative retributivism does not involve the notion that the desert of offenders supplies a reason to think that punishment constitutes an intrinsic good of some sort, then in my view, we should reject the claim that such an account constitutes a version of negative \emph{retributivism}. Perhaps “side-constrained consequentialism” would be a more apt description of such an account.\footnote{For a famous example of side-constrained consequentialism, see Hart, “Prolegomenon to the Principles of Punishment,” p.9. Cited above note 35. For Hart’s view described (in my view in an important sense incorrectly) as a version of negative retributivism, see Duff, “Penal Communications,” p.7. Again, much hinges upon a conceptual debate, here, concerning when it is appropriate for an account to be deemed “retributivist,” even in a thin, negative sense. For the sake of brevity, I shall set this conceptual debate aside for now.} In the following section, I leave aside negative retributivism and instead focus on positive versions of retributivism in order to investigate the potential merits of the desert of offenders in providing a sufficient justification for criminal punishment.\footnote{One way to describe the distinction between positive and negative versions of retributivism is to highlight that, according to the former, the desert of offenders is sufficient for justifiable punishment, whereas according to the latter, the desert of offenders is merely necessary for justifiable punishment. See Duff, \emph{Punishment}, p.19.} I return to address a version of negative retributivism in section 5.

\section*{3.3 – Retributivist themes}

Important beyond the various controversies surrounding the moral and practical credibility of the desert of offenders as such, a further problem afflicts traditional retributivist accounts of the appropriate justification for criminal punishment: defenders of retributivism must explain, in some plausible way, the

\begin{footnotesize}
\begin{enumerate}
\item \footnote{Ibid (Moore’s italics).}
\item \footnote{For a famous example of side-constrained consequentialism, see Hart, “Prolegomenon to the Principles of Punishment,” p.9. Cited above note 35. For Hart’s view described (in my view in an important sense incorrectly) as a version of negative retributivism, see Duff, “Penal Communications,” p.7. Again, much hinges upon a conceptual debate, here, concerning when it is appropriate for an account to be deemed “retributivist,” even in a thin, negative sense. For the sake of brevity, I shall set this conceptual debate aside for now.}
\item \footnote{One way to describe the distinction between positive and negative versions of retributivism is to highlight that, according to the former, the desert of offenders is sufficient for justifiable punishment, whereas according to the latter, the desert of offenders is merely necessary for justifiable punishment. See Duff, \emph{Punishment}, p.19.}
\end{enumerate}
\end{footnotesize}
relationship between the desert of offenders and the justification for inflicting hard treatment upon them. In other words, retributivists must explain precisely what wrongdoers supposedly deserve to suffer, and why. Without a compelling argument establishing this relationship, retributivism looks problematically barbaric.

The problem, here, is that what it means for wrongdoers to “deserve to suffer” punishment is open to several possible interpretations, all of which are, in one way or another, open to objection. It is worth briefly explaining and addressing some of the possible interpretations of retributivism found in the literature.

3.3.1 – Circular retributivism

One possibility is that what it means for wrongdoers to deserve to suffer punishment is that there exists a ready-made justification in place for punishing those who commit serious wrongs. But this interpretation is clearly unacceptable. It would beg the question to hold that punishment is justifiable simply because wrongdoers deserve to suffer punishment, if what we mean by “deserve to suffer punishment” is simply that “punishment is justifiable.” Such a view would be viciously circular, and so unhelpful in facing up to the objection that retributivists must explain, in some plausible way, why wrongdoers deserve to suffer punishment.

---

60 See Duff, “Penal Communications,” p.8.
3.3.2 – Hedonic-suffering retributivism

Another possibility is that what it means for wrongdoers to deserve punishment is that it is right or just that wrongdoers suffer the hedonic pain of punishment, in a proportional although non-equivalent relation, presumably, to the pain and suffering meted out via the commission of their crimes.\(^{62}\) This argument avoids the circularity objection. However, the idea that wrongdoers deserve to suffer a proportional degree of hedonic pain seems implausible on other grounds.

The first problem has to do with the idea of desert. One might ask, precisely what does it mean for someone to deserve a certain form of treatment? One possibility is that the laws of society and the institutions that uphold them provide content to the claim that wrongdoers deserve to suffer punishment. But an institutional construal of desert is unhelpful in this context, since it would merely draw our attention to a set of empirical facts (i.e. that there is in fact a set of rules that make it the case that wrongdoers deserve to suffer punishment) and not any substantive norm (i.e. that wrongdoers for some reason ought to suffer punishment). Another possibility is that there exists an extra-legal or supra-institutional construal of desert that can provide content to the assertion that wrongdoers deserve to suffer the hedonic pain of punishment.\(^{63}\) The problem

---

\(^{62}\) For example see Davis, L., “They Deserve to Suffer,” *Analysis* 32:4 (1972). I ignore in what follows the *lex talionis* interpretation of this view (i.e. an eye for an eye, a tooth for a tooth, a life for a life), since I take it as an obvious and unproblematic point that punitive torture and punitive rape are always unjust under any conditions.

\(^{63}\) For a version of this point, see Feinberg, J., "Justice and Personal Desert," in *What Do We Deserve?: A Reader on Justice and Desert* ed. Louis Pojman &
with this strategy is that such an account seems objectionably vague. It is unclear, on such a view, how it would be possible to account for what forms of punitive treatment wrongdoers supposedly deserve to suffer, and why.\(^6^4\)

The second problem with this view has to do with the idea of suffering. Hedonic-suffering retributivism, it seems to me, is vulnerable to the following objection: if the idea that wrongdoers deserve to suffer a proportional degree of hedonic pain is sufficient for justifiable punishment, then as an outcome of such a view, any form of punishment, insofar as it counts as a way of inflicting pain, might in principle be utilized in a way that purports to be right or just. But this outcome is absurd. Many brutal or dehumanizing forms of punishment are always unjust, no matter how grievous the offences of any given wrongdoer, and no matter how carefully each instance of punishment is designed to generate the appropriately proportionate degree of hedonic suffering, e.g. extreme forms of corporal punishment, punitive bodily mutilation, etc.

---

\(^6^4\) See Feinberg, J., “The Expressive Function of Punishment” *The Monist* 49:3 (1965) pp.421-3. Also see Duff, “Penal Communications,” p.8. Consider the following example: Some defenders of retributivism advocate capital punishment. For example see Kant, *The Metaphysics of Morals*, cited above note 47. Others, however, do not, or at least are unsure whether retributivism justifies an institution of capital punishment. For example see Nozick, *Philosophical Explanations*, p. 378. On what grounds according to this view, one might ask, are we to establish whether capital punishment ought to be considered as among the forms of punishment that wrongdoers putatively deserve, in the hedonic-suffering sense? It seems to me that, on this view, nothing other than bare intuition is able to do the normative work, here, and so such an account seems objectionably vague in its prescriptions concerning what forms of punishment wrongdoers supposedly deserve to suffer, and why.
In my view, this problem relates to a more basic problem with certain traditional versions of positive retributivism, namely, that certain versions of retributivism present a merely formal thesis concerning justifiable punishment: according to some of its variations, retributivism stipulates nothing more than a formal, justificatory relationship between two elements, namely, the grievousness of the wrongdoing in question, and the severity of the suffering meted out via punishment. The view is therefore left incomplete, and so must be supplemented with substantive normative considerations that provide more content to terms such as “desert” and “suffering.” I now turn my attention to three retributivist accounts that attempt to do just that.65

3.3.3 – Forfeited-rights retributivism

In his classic work “Persons and Punishment,” Herbert Morris defends a version of “forfeited-rights” retributivism. According to Morris,

If a person fails to exercise self-restraint even though he might have and gives in to … inclinations, he renounces a burden which others have voluntarily assumed and thus gains an advantage which others, who have restrained themselves, do not possess. This system, then, is one in which the rules establish a mutuality of benefit and burden

65 Importantly, in my view, all three of the following accounts have something attractive in their respective justifications for inflicting hard treatment upon wrongdoers. I return to this point in the final section (section 5). It is in part for this reason that my comments should not be taken as providing knockdown arguments against these versions of retributivism. However, in my view, as will be discussed in what follows, it is possible to capture certain attractive ingredients associated with these variations of retributivism without appealing to the desert of offenders as such. This point will be discussed more extensively in Chapters 4 and 5.
and in which the benefits of noninterference are conditional upon the assumption of burdens.  

For Morris, therefore, if a person fails to assume the burdens of compliance with the criminal law, in doing so she enjoys an unfair advantage over the wider, law-abiding community. As a result, she thereby forfeits, or in Morris’s view she ought to forfeit, the benefit of the protections afforded by such a system that are, he argues, “conditional upon the assumption of burdens.”  

Such an account is a retributive theory insofar as it provides a strictly backward-looking account of what justifies criminal punishment, as well as a fuller account of why it ought to be considered right or just in itself that wrongdoers are made to suffer punishment.  

In my view, Morris is correct that reciprocity about basic rights ought to factor centrally into our thinking about what justifies, normatively speaking, the practice of criminal punishment. Also, there is something attractive (but nevertheless contestable) about the idea that wrongdoers gain an unfair advantage over the wider, law-abiding community. Morris’ argument in this context is structurally similar to the complaint against so-called “free-riders” in egalitarian theories of distributive justice: it is unfair if people who benefit from a

---

67 Ibid.
69 This point will be discussed further in Chapter 5.
set of institutions fail to do their part in shouldering the burdens associated with such institutions.

However, the unfair advantage theory of forfeited-rights advanced by Morris does not hold up under closer scrutiny. The basic problem is that Morris’s view presupposes an untenable account of what is essential to criminality, i.e. the relevant aspect of crime meriting a response on the part of the normative community. Commentators have observed that in no meaningful sense do torturers or rapists, for example, gain an unfair advantage over the wider, law-abiding community. For almost everyone, there is no burden involved in refraining from torturing or raping others, and so no advantage of the sort Morris describes accrues to those who violate others, or on Morris’ account, “fail to exercise self-restraint…” More basically, it would be objectionable to construe the majority of law-abiding citizens as “exercising self-restraint” in the face of the supposed incentive to harm others. The unfair-advantage theory of forfeited rights therefore fails, since many forms of criminal wrongdoing do not involve the sort of unfair advantage Morris specifies – i.e. the so-called advantage of not bothering to shoulder the burden of restraint. The view that what wrongdoers deserve to suffer is the forfeit of the unfair-advantage that attaches to their noncompliance is thus ultimately untenable.

---

71 For a similar point see von Hirsch, *Censure and Sanctions*, p.8. Although von Hirsch argues, “… it is straining to try to explain the heinousness of common crimes such as burglary and robbery…” with reference to the unfair-advantage that supposedly accrues to wrongdoers.

3.3.4 – Punitive emotions

Some retributivists aim to establish the relationship between punishment and the desert of offenders by appealing to the punitive emotions that ought to be considered the intrinsically appropriate response to crime. Put in rather harsh terms, the classical punishment theorist J.F. Stephen argues that criminal punishment is the appropriate expression of the “hatred and … vengeance” that criminal wrongdoing “excites in healthily constituted minds.”

More recently, Moore argues that moral guilt is the intrinsically appropriate response to crime on the part of the offender, and that therefore, if you or I were guilty of an egregious form of criminal wrongdoing, the extreme extent of our own guilt would be such that we would undoubtedly endorse a punitive response to our own criminal wrongdoing, i.e. we would rightly judge that we ourselves deserve to suffer. Thus, if we are to treat wrongdoers with the respect they are owed as moral decision-makers such as ourselves, i.e. as rational and responsible moral agents, then we should also endorse a punitive response to criminal wrongdoing in the general case. He argues,

…we should ask ourselves what …[a criminal wrongdoer] deserves by asking what we would deserve had we done such an act. In answering this question we should listen to our guilt feelings, feelings whose epistemic import is not in question in the same way as … [some other feelings]. Such guilt feelings should tell us that to … [commit an egregious criminal act] is to forfeit forever any lighthearted idea of going on as before. One should feel so awful that the idea of again leading a life unchanged from before, with the same goals and hopes and happiness, should appear revoltingly incomprehensible.

74 Moore, *Placing Blame*, p.149.
75 *Ibid* (Moore’s italics).
For Moore, guilt is a “virtuous” emotion, in that virtuous individuals who seriously wrong others feel an appropriate degree of guilt. Therefore, we should trust that guilt, as a virtuous emotion, is a good heuristic guide to what morality requires as a response to criminal forms of wrongdoing on the part of ourselves, and thus on the part of wrongdoers generally.

Moore’s view, I submit, has many appealing characteristics. In my view he is right to point out that guilt (understood in terms of a profound feeling of remorse) is an aspect of the appropriate response to crime on the part of wrongdoers (although, importantly, such a claim is not entirely uncontroversial). However, the main issue with Moore’s account, I submit, is that he relies heavily on an undefended intuition, namely, that morally appropriate guilt implies, in a straightforward way, the view that wrongdoers deserve to suffer punishment. There is a significant unaccounted for gap, so to speak, between the view that wrongdoers ought to feel profoundly guilty, and the view that wrongdoers, because they ought to feel profoundly guilty, deserve to suffer punishment.76

However, Moore’s account is not without merit. While his account involves certain weaknesses surrounding his lengthy and complex attempt to ground a justification for an institution of criminal punishment in the desert of offenders (with reference to the virtuous emotion of guilt), his argument that a certain moral

---

76 Duff, *Punishment*, p.25. As Duff eloquently argues, “…despite its detail and complexity, Moore’s argument appears to amount to little more than an appeal to the intuition (expressed in first-person cases through the emotion of guilt) that “the guilty deserve to suffer”: it does not tell us why they should suffer, or why guilt should generate the judgment that I ought to suffer, or what they ought to suffer, or why it should be a proper task for the state to inflict that suffering.”
response (that includes an emotional dimension) on the part of offenders ought to be considered morally required seems plausible. I return to this point in Chapter 5.

3.3.5 – Expressive retributivism

Expressive retributivists claim that the wrongdoer has flouted correct values, and hence ought to suffer punishment as the intrinsically appropriate response to crime on the part of the community, in order to express to her and others the importance of correct values. Robert Nozick argues that we ought to view the value of retributive punishment "nonteleologically, so that it is seen as right or good in itself..."  More specifically, for Nozick, “The wrongdoer has become disconnected from correct values, and the purpose of punishment is to (re)connect him. It is not that this connection is a desired further effect of punishment: the act of retributive punishment itself effects this connection.”

Again, this sort of view has many appealing aspects. Given egregious instances of criminal wrongdoing, a condemnatory reaction of some sort on the part of those who have been wronged (including members of the normative community who have not themselves been directly victimized) seems, at least in most cases, appropriate, and seems essential to any defensible account of

---

77 Nozick, Philosophical Explanations, p.374.
78 Ibid. Another way of putting the point is that, for Nozick, the wrongdoer “flouts correct values,” and the purpose of punishment is to re-connect her with correct values “in a second best way.” See Nozick, Philosophical Explanations, p.382.
morality. However, the central problem with expressive punishment on the account defended by Nozick is that it need not address the wrongdoer as a rational and responsible moral agent, in the sense that, on Nozick’s view, the value of expressive punishment is established without reference to how the moral message is received or interpreted by the wrongdoer. The wrongdoer is simply thrashed, so to speak, in the name of correct values. Furthermore, a related objection is that Nozick’s view seems to conflate (I) the expression of blame or censure, with (II) the practice of punishing wrongdoers, in the sense that his account seems to fail to acknowledge that (I) might be accomplished in a variety of ways that do not involve (II) per se.

In my view, the essential problem with traditional forms of retributivism discussed up to this point is their inability to explain in a compelling way the idea that wrongdoers “deserve to suffer” punishment. The idea is either left objectionably vague, or else is explained in a way that is, in one way or another, substantively objectionable. However, the contestable thought that wrongdoers deserve to suffer punishment expresses a deeper and more plausible moral

---

79 However, in my view, we should distinguish more carefully between the condemnatory judgment that an instance of wrongdoing makes appropriate, and the appropriate mode of expressing such a condemnatory judgement. This point will be taken up further in Chapter 4.
80 See Duff, “Penal Communications,” pp.32-3. As Duff argues, “…we should distinguish between expression and communication, and see punishment as communicative rather than expressive…” since communication “…addresses the other as a rational agent: expression need not address her at all, and might not address her as a rational agent.” Duff’s view will be explained in further detail in what follows.
Consider the following moral claim: there is something seriously morally amiss – i.e. a moral wrong occurs – when those who perpetrate egregious moral wrongs proceed (and are allowed to proceed) to live their lives as if no wrongdoing had occurred at all. Nonetheless, in my view, the thought that wrongdoers deserve to suffer punishment fails to appropriately capture our firmly held intuition that it is unjust and unfair when egregious instances of wrongdoing go unrecognized and unaddressed.

If a version of the desert-based view is left objectionably vague, then it remains unclear how to circumscribe the scope of the basic rights of offenders, since it remains unclear precisely what wrongdoers supposedly deserve to suffer, and why. If, on the other hand, the desert principle is explained in a way that is substantively objectionable, then we have reason to take issue with how the basic rights of wrongdoers are circumscribed according to such a view. Before moving on, however, it is important not to lose sight of the most valuable aspect of retributivism: the point of departure for retributivists is to attempt, in the context of criminal justice, to treat wrongdoers with the respect they are owed as rational and responsible moral agents. In my view, any plausible theory of punishment

---

82 In my view, however, the appropriate backward-looking response to wrongdoing is not best articulated in terms of desert, but instead in terms of other familiar moral ideas, such as responsibility, accountability and blame. This point will be discussed further in Chapter 4.

83 What might be behind the contestable retributivist notion that wrongdoers deserve to suffer punishment is the plausible notion that wrongdoers ought to be held accountable in some way, and that victims have a right to be recognized as victims, i.e., as morally considerable individuals who have legitimate claims to having been seriously wronged by others. For a similar point, see Scanlon’s discussion of the importance of “affirmation” in the context of criminal justice. See Scanlon, “Punishment and the Rule of Law,” p.221.

84 This point will be taken up further in Chapters 4 and 5.
must attempt to do the same. Nonetheless, if we are not presented with a clear and compelling account of what wrongdoers supposedly deserve to suffer, and why, then retributivism looks problematically barbaric. Furthermore, absent a compelling argument that establishes the relationship between the desert of offenders and justifiable punishment, we are unlikely to be convinced that we ought to set up a set of coercive institutions in order to mete out the punishment that wrongdoers supposedly deserve to suffer.

3.4 – Retributivist compromise theories explored

To briefly sum things up so far: strict consequentialism about punishment, whether sophisticated or simplistic in its articulation of the appropriate forward-looking aims of criminal justice institutions, tends to fail to adequately account for the moral status of offenders as rational and responsible moral agents, and as individuals who possess basic rights against being manipulated or used “as a means” in certain ways. That having been said, strict consequentialism is commendable in its emphasis on the importance of justifying punishment with respect to tangible, forward-looking benefits. Punishing wrongdoers is a highly costly enterprise, both morally and otherwise, and so justifying a public institution of criminal punishment seems more plausible from the perspective of trying to ascertain what ought to be considered the empirically verifiable valuable outputs of such an institution, beyond merely “doing justice.”

Traditional forms of retributivism, on the other hand, are commendable in that they take as their point of departure an effort to treat wrongdoers in ways
that respect their status as moral agents. Furthermore, such accounts seek to
ground the liability of wrongdoers to suffer punishment, and hence an account of
their basic rights in the context of criminal justice, in a concrete moral principle,
namely, the principle of desert (however understood). The issue with
retributivism, in my view, is precisely that the principle of desert is highly
controversial – respectively, the various articulations of how the desert of
offenders relates to justifiable punishment are either left objectionably vague, or
else are explained in objectionable ways.

As discussed in Chapter 2, in an effort to recognize and address the
various pitfalls associated with both strictly consequentialist accounts and
traditional, positive versions of retributivism, more recently, a third approach has
emerged: compromise theories attempt to bridge the gap between exclusively
forward-looking and exclusively backward-looking accounts, typically with the aim
of reconciling in some way retributivism and consequentialism. As we have seen,
generally speaking, compromise theories involve attempting to justify the
punishment of wrongdoers with reference to forward-looking aims, but aim to limit
or constrain the pursuit of such aims with reference to some form of non-
consequentialist desiderata for permissible punishment. In the case of
compromise theories that aim to reconcile retributivism and consequentialism, the
non-consequentialist principle underlying the liability of wrongdoers to suffer

---

85 In contrast with Hart’s non-retributivist compromise theory addressed in
Chapter 2.
punishment (and thus, to be “used” for the sake of achieving the forward-looking aims of criminal justice institutions) is the now familiar notion of desert.\(^{86}\)

**3.4.1 – Punishment as public censure plus prudential supplements**

Andrew von Hirsch defends a theory of punishment that involves an attempt to reconcile desert theory with the aim of crime prevention (via a system of punishment that focuses on general deterrence). He argues that the expression of public censure ought to be considered the intrinsically appropriate response to criminal forms of wrongdoing. In this sense, he defends a version of retributivism about punishment. However, von Hirsch also finds a role in his account for general deterrence, i.e. the forward-looking aim of attempting to prevent crime via the use of deterrent criminal sanctions. Importantly, on this view, the pursuit of crime prevention via general deterrence ought to be constrained with reference to the desert-based, censuring aspect of criminal punishment. This, for von Hirsch, is the aspect of the criminal sanction that addresses wrongdoers as moral agents. Therefore, we can construe this view as a version of negative retributivism, in that the pursuit of the forward-looking aim of crime prevention ought to be constrained with reference to the desert of

---

\(^{86}\) It is important to note that the view I aim to defend in what follows (particularly in Chapters 4, 5 and 6) is itself a compromise theory, and so the views that I describe in the following section, according to my argument, come close (i.e. closer than more traditional accounts) to acceptable theories of punishment. For my purposes, it is worth reviewing two famous retributivist compromise theories in order to point out some of the strengths and weaknesses of presently existing accounts, and, importantly for my purposes, to help situate in context the non-retributivist restorative justice compromise theory I aim to articulate and defend in the remainder of the thesis.
offenders. In this way, von Hirsch attempts to avoid the charge that, by pursuing crime prevention via general deterrence, he is treating wrongdoers as something other than fully rational and responsible moral agents. More specifically, von Hirsch argues that public censure is not the only constitutive element of criminal punishment that must be justified. A theory of punishment, he argues, must also account for the “hard treatment” that is essentially involved in the practice of punishing wrongdoers. He therefore takes seriously the famous complaint regarding expressive versions of retributivism, e.g. Nozick’s version of expressive retributivism, namely, that a public expression of condemnation need not take the form of criminal punishment per se.

This view is therefore a retributive theory of punishment whereby the central purpose of punishing wrongdoers is two-fold: (1) to censure wrongdoers for their wrongful conduct, and (2) to provide a “supplementary prudential disincentive” to wrongful conduct via penal hard treatment. Regarding (1), von Hirsch argues, “The criminal law, through the censure embodied in its prescribed sanctions, conveys that the conduct is wrong, and the moral agent thus is given [strictly moral] grounds for desistence.” The censuring aspect of the criminal sanction, he makes clear, can be dissociated from the hard treatment aspect, in that mere formal censure – e.g. a public admonishment of some sort – without the addition of penal hard treatment is a possible response to wrongdoing on the part of the state (although mere formal censure without hard treatment would not

---

constitute criminal “punishment” per se, since it would lack one of the two constitutive elements). He writes, “A condemnatory response to injurious conduct … can be expressed either in a purely (or primarily) symbolic mode; or else, in one in which the reprobation is expressed through the visitation of hard treatment. The criminal sanction is a response of the latter kind.”

Due to the moral fallibility of human beings, von Hirsch argues, the censuring aspect of the criminal sanction ought to be supplemented by a prudential disincentive in the form of penal hard treatment. In this way, we are given a reason going beyond the censuring role of punishment, that is the strictly speaking moral reasons not to offend. Thus, regarding (2), von Hirsch argues, “The preventative function of the sanction should be seen … as supplying a prudential reason that is tied to, and supplements, the normative reason conveyed by penal censure.” Importantly, for von Hirsch, the purpose of the censuring aspect of the criminal sanction is not to reform the offender, or to attempt to evoke in her certain sentiments, such as guilt, remorse or regret. He writes, “Censure gives the actor the opportunity for … responding [in the appropriate way], but it is not a technique for evoking specified sentiments.”

According to von Hirsch, it is not the appropriate role of the state to pursue moral reform on the part of offenders, although criminal sanctions should provide, on his view, the opportunity for moral reform.

---

92 *Ibid*.
94 *Ibid*. This point will be discussed further in Chapter 6.
One further integral aspect of von Hirsch’s account is worth highlighting before proceeding: the censuring aspect of the criminal sanction has primacy over the preventative aspect, in order to preserve the sense in which penal censure addresses wrongdoers as rational and responsible moral agents. Furthermore, he argues, the censuring aspect serves to anchor the proportionality principle underlying his account. He argues that the censuring aspect of the criminal sanction provides grounds for both ordinal and cardinal proportionality in punishment: it implies ordinal proportionality, in that the greater the seriousness of the wrongdoing in question, the more blameworthy the offender, and thus the greater the severity of censure that is warranted as a response. Interestingly, von Hirsch also argues that the censuring aspect implies cardinal proportionality: if penal hard treatment is to provide merely a prudential supplement to, rather than a replacement for, the normative reasons not to offend inherent in the criminal sanction, then the severity of the sanction must not be so overwhelming that it supplants or silences the normative reasons not to offend. As such, von Hirsch defends a “decremental strategy,” whereby criminal sanctions ought to be scaled down over time, with the aim of eventually imposing no more than three year sentences of imprisonment for serious criminal offences, with the exception of homicide, which he argues should carry no more than a five year sentence of imprisonment. In this way, he aims to avoid the charge that, by including a role for the pursuit of general deterrence, his account fails to treat

---

wrongdoers with the respect they are owed as rational and responsible moral agents.

There are many appealing aspects associated with this view. Importantly, von Hirsch’s account involves an attempt to reconcile the forward-looking aim of crime prevention with the backward-looking aim of public censure. This view therefore attempts to treat wrongdoers in a way that demonstrates respect for their status as rational and responsible moral agents, while also taking seriously the notion that criminal punishment ought to pursue beneficial forward-looking aims, e.g. crime prevention. Along these lines, by construing the aim of crime prevention as a supplement to (rather than a replacement for) the moral reasons not to offend, von Hirsch emphasizes the importance of not treating the wrongdoer as a mere means toward the end of crime prevention. His account thus aims to treat wrongdoers as moral agents, although less than perfect moral agents, that can be motivated both by moral as well as prudential reasons not to offend.

However, by identifying crime prevention via general deterrence as the appropriate forward-looking aim of criminal justice institutions, there remains a deep, unresolved tension between the specified forward-looking and backward-looking aspects of this view. It is important to question whether or not, for example, five years imprisonment for the crime of homicide can plausibly be understood as a supplement to the moral voice of the criminal law. As Goldman argues, a paradox emerges in any attempt to reconcile retributivism with

---

96 For this point see Duff, “Penal Communications,” p.44.
deterrence theory: punishments that are constrained by a principle of desert are unlikely to be sufficiently weighty from the perspective of deterrence, and so unlikely to deter anyone with any degree of success. On the other hand, punishments that are likely to successfully deter would be ruled out by the constraints imposed by any reasonable construal of the desert of offenders. This sort of complaint is relevant to von Hirsch’s suggestion that criminal punishments should be scaled down, such that homicide, for example, carries no more than a five-year sentence of imprisonment. In an effort to rein in the deterrent effect of the law (so as to conform to the moral requirement that it is unacceptable to treat wrongdoers as less than fully rational and responsible moral agents), we are left with an approach to criminal punishment that is unlikely to deter anyone.

On the other hand, were we to hold on to von Hirsch’s general account, but ignore his more particular suggestion that we ought to scale down the anchor for cardinal proportionality, (say to include, as many systems of criminal punishment in fact do, fifteen to twenty-five year prison sentences for homicide, etc.), then it is no longer apparent that hard treatment is providing merely a prudential supplement to the aim of public censure and the moral reasons not to

---


98 I owe this point to Duff. See Duff, “Penal Communications,” p.44.

99 For this point see Duff, “Penal Communications,” p.45. Importantly, in agreement with Duff, I am not suggesting here that criminal punishment institutions ought to aim at general deterrence. Rather, I aim attempting to highlight an unresolved tension inherent in von Hirsch’s account as stated.
offend; on this revised version of the view, the prudential reasons not to offend would completely depose the moral voice of the criminal law.\textsuperscript{100}

Also, we should question the implications of the desert principle underlying the suggestion that penal hard treatment ought to be understood in terms of a prudential supplement. Is it acceptable, one might ask, to construe hard treatment as a supplement to what wrongdoers deserve, understood in terms of merely formal, public censure? Many might argue that formal public censure is too thin an account of what wrongdoers deserve to suffer, and thus to construe, for example, five years imprisonment for homicide as a supplement to the moral voice of the criminal law is, in an important sense, wrongheaded. We might plausibly think that, despite von Hirsch’s suggestion that hard treatment serves no other purpose than as a prudential supplement, the burdens involved in the practice of punishing wrongdoers are an aspect of what wrongdoers, morally speaking, owe to others (i.e. their victims, their victims families, the normative community, etc.), as opposed to something wrongdoers ought to suffer in addition to public censure, solely for the sake of providing others with prudential reasons not to offend.

In short, on this account there seems to be a deep, unresolved tension between the aims of (1) meting out the censure that wrongdoers deserve to suffer, on the one hand, and (2) providing prudential supplements to deter crime, on the other. Furthermore, by attempting to reconcile the desert of offenders with the aim of preventing crime via general deterrence, we are left with an

\textsuperscript{100} See Duff, “Penal Communications,” p.44-45.
implausibly thin account of the desert of offenders. In many cases at least, it seems plausible that wrongdoers owe something more to those whom they have wronged, morally speaking, than merely suffering a form of formal, public censure per se.\textsuperscript{101} It seems wrongheaded, therefore, to construe the burdens of criminal punishment (i.e. the “hard treatment” involved in criminal punishment) in terms of a prudential supplement to the moral voice of the criminal law. Morally speaking, wrongdoers arguably owe something more to their victims and the normative community.\textsuperscript{102}

3.4.2 – Communicative punishment, or “teleological” retributivism

This leads us to a different, more normatively ambitious retributive compromise theory, which construes criminal punishment as a “penitential process,”\textsuperscript{103} designed both to constitute a mode of forceful public censure, and to bring the wrongdoer to reform herself, repent her crime(s), and thereby reconcile herself with those whom she has wronged. According to R.A. Duff’s “communicative theory of punishment,” criminal punishment ought to communicate the blame or censure that wrongdoers deserve. The censuring aspect of communicative punishment, for Duff, is essentially retributivist in character: he writes, “…we should use penal hard treatment to communicate the censure that offenders deserve…”\textsuperscript{104} However, for Duff, in order to be justifiable, the communication of deserved censure must be conducive to fostering moral

\textsuperscript{101} Cf. Duff, \textit{Punishment}.\textsuperscript{102} This point will be taken up further in Chapters 4 and 5.\textsuperscript{103} See Duff, “Penal Communications,” p.88.\textsuperscript{104} See Duff, \textit{Punishment}, p.30.
recognition in wrongdoers, toward the further ends of reparations between wrongdoers, their victims, and the wider, normative community. We may refer to these ends collectively – i.e. the ends of fostering recognition in wrongdoers, and reparations between wrongdoers, their victims, and the normative community – as the ends of restorative justice. Duff writes,

Communicative punishment … should be understood and justified as a communicative, penitential process that aims to persuade offenders to recognize and repent the wrongs they have done, to reform themselves, and so to reconcile themselves with those they have wronged.

Therefore, on this view, punishment is justified in part with reference to its contribution to pursuing the forward-looking ends of restorative justice, rather than, as on von Hirsch’s account, crime prevention via general deterrence. However, like von Hirsch, Duff firmly resists the idea that punishment ought to be considered valuable purely as a means to achieving further, valuable ends (i.e. in the case of Duff’s view, the ends of restorative justice). Retributive punishment, he argues, ought to play a necessary or constitutive role in our pursuit of the ends of restorative justice. He writes that

To take wrongs seriously as wrongs involves responding to them with criticism and censure … This is not to say that we should censure a wrongdoer only when we believe that there is some chance of … persuading him. We may think that we owe it to his victim, to the values he has flouted, and even to him, to censure his wrongdoing.

---

105 Of course, there is widespread disagreement in the literature concerning how to understand what constitutes “restorative justice.” The restorative justice alternative approach to the problem of punishment will be discussed in greater detail in Chapter 4.
even if we are sure that he will be unmoved and unpersuaded by the
censure. But our censure still takes the form of an attempt (albeit what
we believe is a futile attempt) to persuade him.  

Importantly, Duff views communicative punishment as first and foremost a form
of public censure: criminal punishment metes out the censure that wrongdoers
deserve. However, if criminal punishment is to be a morally defensible form of
public censure, it must also seek restorative justice. On Duff’s view, therefore,
the punishment of wrongdoers is valuable as a means to achieving the ends of
restorative justice, but the value of punishing wrongdoers is not *purely* derivative
from such ends. The punishment (understood in terms of a mode of forceful
public censure) of wrongdoers is the appropriate response to crime on the part of
the normative community, but only insofar as it takes meaningful aim at the ends
of restorative justice. The upshot is that, for Duff, criminal punishment may retain
its value even in cases wherein it fails to achieve, *and it is known that it will fail to
achieve*, its desired goals. Making his view clear, he writes, “Our responses to
crime should aim for ‘restoration,’ for ‘restorative justice’: but the kind of
restoration that criminal wrongdoing makes necessary is properly achieved
through a process of retributive punishment.”  

An element of retributivism is
therefore essential to Duff’s account.

But, one might ask, is Duff’s account as stated a version of positive or
negative retributivism about punishment? It seems to me that Duff’s account
involves elements of both. For Duff, public censure in the form of criminal
punishment (conditioned by the pursuit of the ends of restorative justice) is the

---

appropriate backward-looking response to criminal wrongdoing. We are thus provided with a positive, backward-looking reason to punish wrongdoers. However, the pursuit of the ends of restorative justice, in order to be permissible, ought to be constrained according certain non-consequentialist principles, including the retributive principle that communicative punishment must be the mode of forceful public censure utilized for the sake of pursuing the ends of restorative justice, as well as the principle that communicative punishment must always address the wrongdoer as a rational and responsible moral agent.

Therefore, it is possible to characterize Duff’s view as a form of teleological communicative retributivism about punishment. The teleological retributivist holds that retributive punishment is the appropriate backward-looking response to crime, so long as it conforms to a reasonable attempt to achieve certain ends or goals. However, on such a view, the permissibility of punishment is established independently of the actual achievement of such ends. Therefore, the forward-looking dimension of Duff’s account (i.e. the emphasis on the ends of restorative justice) applies, in a sense, as a check or restriction on the positive, backward-looking reason to punish wrongdoers, i.e. retributive punishment is valuable only insofar as it aims at the ends of restorative justice. On the other hand, the backward-looking dimension of Duff’s account (i.e. that criminal punishment metes out the censure that wrongdoers deserve) applies as a check or restriction on the appropriate means that may be employed in order to pursue its forward-looking aims, i.e. the ends of restorative justice.

For the characterization of this sort of view as “teleological” retributivism, see Nozick, *Philosophical Explanations*, pp.374-5.
Duff’s contention that his view involves a retributive dimension has given rise to various criticisms of his account. For example, Honderich argues that

… punishment on this view does not have to have even a chance of reforming in order to be justified. ... What the view comes to ... is that punishment is right when it tells an offender something ... even if it [i.e. the punishment] and his understanding [of it] have no reformative effect on him. The thought may incline us to suppose this is more of a retribution theory than so far supposed, purer retribution than supposed.111

More straightforwardly, Honderich writes that Duff’s communicative account “...is at bottom a traditional retribution theory, most likely about culpability and distress in some way... What goes together with the proposition that punishment is communication is just that punishment is deserved.”112

Duff’s account may be described as one view among a cluster of possible views that understand punishment as valuable for the sake of pursuing the ends of restorative justice, although in Duff’s case, with a commitment to retributivism as a necessary or constitutive aspect of the appropriate pursuit of restorative justice. Duff's account is unlikely to appeal to those who are skeptical of the moral and practical credibility of the desert principle essential to all retributivist accounts. For Honderich and others, Duff’s view collapses into a form of traditional retributivism when the intrinsic value of punishment (understood as a mode of forceful public censure) is emphasized over and above the actual

achievement of the more ambitious aims associated with the restoration of wrongdoers, and the reconciliation of wrongdoers with their communities.\textsuperscript{113}

Furthermore, Duff’s account seems vulnerable to a separate complaint concerning expressive and communicative forms of retributivism more generally. Like Nozick’s account, (but unlike von Hirsch’s account), commentators have observed that expressive or communicative retributivism has the potential to conflate (I) the expression / communication of blame or censure, with (II) the practice of punishing wrongdoers, in that (I) might, at least in principle, be accomplished in a variety of ways that do not involve (II) per se.\textsuperscript{114} Duff’s view and his rivals will be discussed in further detail in Chapters 4 and 6.

### 3.5 – Some further remarks concerning retributivism

Thus far, I have focused on two related problems concerning desert theory. Traditional retributivists rely on a principle of desert, i.e. that wrongdoers in some sense “deserve to suffer” punishment, which is often either left objectionably vague, or else specified in a way that is substantively objectionable. In its more simplistic forms, desert theory is objectionably indeterminate in its prescriptions concerning what wrongdoers deserve to suffer,

\textsuperscript{113} This observation has inspired others to defend a (restorative justice motivated) version of traditional, expressive retributivism akin to that of Nozick. See Bennett, \textit{The Apology Ritual}, esp. Chapter 7.
\textsuperscript{114} See Feinberg, “The Expressive Function of Punishment,” p. 423. Also see Honderich, \textit{Punishment: The Supposed Justifications}, p.182. Honderich argues, “[A] question may have arisen in your mind before now, one raised by any theory of punishment as communication. If your aim is to \textit{tell somebody something}, is it necessary to do so [for example] by putting him in jail for 20 years?” (Honderich’s italics).
and why. According to its more sophisticated formulations, desert theory provides objectionable (or at least, highly controversial) reasons why we should accept the thesis that the desert of offenders provides a sufficient justification for inflicting hard treatment upon them in all cases (or, at least in the vast majority of cases), and thus why it should be the case that a just society ought to set up a costly set of institutions to mete out the suffering that wrongdoers supposedly deserve. The challenge for retributivists, therefore, is to specify a compelling account of what it means for wrongdoers to deserve to suffer punishment, and why they should be made to suffer by others, e.g. the state. Without a compelling account in hand, it is unclear why suffering is something that should be inflicted upon wrongdoers, why the suffering of wrongdoers ought to be considered an intrinsic good, and furthermore, why it should be considered the appropriate role of the state to mete out the suffering that wrongdoers supposedly deserve.

Of course, there are plausible ideas inherent in many of the different attempts to articulate the content of the principle of desert essential to retributivist accounts. Forfeited-rights retributivism is in certain respects compelling, as it highlights that a system of criminal punishment ought to be related to fairness in civil society. The forfeited-rights account also provides an explanation for why criminal punishment ought to be considered the appropriate role for the state; it is in an important sense unfair if wrongdoers are not held accountable to their victims and the community for their wrongful conduct. However, it seems implausible to suggest that criminal justice concerns nothing other than the equal distribution of fairness in civil society. The idea that all examples of criminal
wrongdoing, and therefore the appropriate response to all examples of criminal wrongdoing, can be reduced to a concern about unfair advantage seems unlikely. Criminal wrongdoing also, at least in many cases, involves a deeper moral problem, i.e. an affront to the moral status of the victim(s), and a failure on the part of offenders to live up to their moral responsibilities to others.

The punitive emotions account is also in certain ways attractive. It seems correct to point out that a certain moral response on the part of offenders ought to be considered morally required. However, one might rightly question whether feelings of profound guilt are all that wrongdoers can be said to owe to their victims and others in the community. Furthermore, one might question whether eliciting the virtuous emotion of guilt grounds a justification for retributive punishment in all cases. Perhaps, at least in certain cases, it is more appropriate that eliciting this emotion grounds another sort of response on the part of the community, given a fuller account of what wrongdoers owe to their victims and others.

In my view, the version of retributivism that comes closest to the restorative account that I defend in the remainder of this thesis is expressive retributivism, i.e. the view that what wrongdoers deserve is censure. However, like punitive emotions, we should question whether or not a responsibility to

---

115 In what follows, however, I argue that we should distinguish more carefully (i.e. more carefully than expressive retributivists typically do) between judgements of blameworthiness, on the one hand, and the expression of such judgements by various possible means, on the other. Also, in contrast with the views defended by deserved censure theorists, I will argue that the desert of offenders plays no basic role (i.e. no role that is not appropriately subject to further elucidation and analysis) in justifying the infliction of hard treatment upon wrongdoers.
censure wrongdoers, or hold wrongdoers in some sense accountable to the community, straightforwardly implies an obligation to inflict hard treatment upon them in all cases. As von Hirsch argues, there are other ways to express censure, e.g. merely formal or symbolic censure, and so the defender of the deserved censure account must provide a further argument concerning why censuring criminal wrongdoers must in all cases take the form of inflicting hard treatment upon them. For von Hirsch, what grounds the hard treatment aspect of the criminal sanction is not censure per se, but rather providing a prudential supplement to the moral voice of the criminal law. Absent the need for a such a prudential supplement, according to von Hirsch, it would be possible for the practice of censuring criminal wrongdoers to take a different form. For Duff, by contrast, the hard treatment involved in the criminal sanction amounts to the censure that is the intrinsically appropriate response to crime. It is what wrongdoers deserve to suffer for their criminal acts. Duff argues that censuring criminal wrongdoers is something that the community owes to victims, to other members of the community, to the values the wrongdoer has flouted, etc.

As will be discussed in Chapter 4, for Duff, censuring criminal wrongdoers is also constitutive of the ends of restorative justice; criminal punishment is a constitutive element of the restorative aim of repairing the normative relationships between wrongdoers and their communities. The purpose of Chapter 4 will be in part to interrogate this claim, and argue that we should reject this retributive component of the communicative theory of punishment.

---

For a different version of this view, see Bennett, *The Apology Ritual*, p.195.
3.6 – Conclusion

There are certain attractive features associated with desert theory. The point of departure for retributivists is to attempt, in the context of criminal justice, to treat wrongdoers and the community as moral agents capable of responding to moral reasons. In this Chapter, however, I argued that desert theory tends to be left objectionably vague, or else tends to be specified in a way that is objectionable. We are now in a position to see, therefore, that both traditional consequentialist accounts and traditional retributivist accounts involve certain deficiencies. In the case of consequentialist accounts, the problem involves failing to treat wrongdoers with appropriate respect, i.e. both in the context of their moral agency and in the context of their basic rights. In the case of retributivist accounts, the problem involves the visitation of hard treatment, i.e. the question of why the community ought to inflict hard treatment upon wrongdoers, or what makes wrongdoers “deserving” of suffering.

In the next Chapter, I explore an approach that has emerged in part out of a deep dissatisfaction with more traditional approaches to criminal justice theory. Defenders of restorative justice alternatives argue that more traditional approaches miss out on important moral values associated with wrongdoing and the appropriate response to wrongdoing, such as the duties of offenders and the values of understanding, recognition, apology and reconciliation. These
alternatives and their relationship to retributivism will be our next subject for
discussion.
Chapter 4 – Moral responsibility and the restorative justice alternative

4.1 – Introduction

In Chapters 2 and 3, I argued that traditional approaches to the problem of punishment remain, at the very least, highly controversial. Strict consequentialism fails to account for the importance of treating wrongdoers with appropriate respect. Criminal justice is about more than simply good outcomes; it is also about appropriate modes of treatment that involve respect for the status of wrongdoers as moral agents who possess basic rights. Traditional versions of retributivism raise a different set of concerns, including the problem of why wrongdoers deserve hard treatment, and why the imposition of hard treatment is necessary in order to treat wrongdoers with the respect they are owed in the context of criminal justice. While various retributivist theories of punishment attempt to establish why wrongdoers deserve to suffer punishment, such accounts often fail to be convincing. We are often left with unanswered questions concerning why the imposition of suffering is a necessary condition for holding wrongdoers accountable to others.

In this Chapter, I explore an alternative approach to the problem of criminal wrongdoing: restorative justice, in which healing or repairing the harm of criminal wrongdoing is emphasized over and above the traditional aims of taking retribution, or punishing for the sake of bringing about good outcomes. Two opposing views dominate the restorative justice literature: defenders of
restorative justice typically adopt either the view that (I) restorative justice stands in stark opposition to retributive justice, and therefore stands at odds with traditional approaches that focus on state punishment as opposed to moral reparation,\textsuperscript{117} or the view that (II) retributive justice is a necessary condition for the pursuit of moral reparation, and therefore restorative justice requires a process of retributive punishment for its realization.\textsuperscript{118}

In this Chapter, I argue that the two dominant views represent a false dichotomy. While both general accounts have certain strengths and weaknesses, ultimately we should reject both views. I argue that we should instead view the relationship between restorative justice and criminal punishment in such a way that punishing wrongdoers is not necessary for the sake of pursuing the ends of restorative justice, but also not necessarily at odds with pursuing such ends.\textsuperscript{119}

What is necessary for the sake of pursuing the ends of restorative justice, I argue, is holding wrongdoers in an appropriate way accountable to their victims and the community. However, by appealing to a non-retributivist theory of moral responsibility and blame, I argue that the particular mode of holding wrongdoers


\textsuperscript{118}See, for example, Duff, “Restorative Punishment,” p. 367. Also see Daly, K., “Restorative Justice: The Real Story,” \textit{Punishment and Society} 4:1 (2002).

\textsuperscript{119}As discussed in Chapter 3, for my purposes, I refer to the aims of (1) \textit{fostering moral recognition in wrongdoers}, and (2) \textit{restoring the normative relationship between wrongdoers, their victims, and the normative community} collectively as “the ends of restorative justice.”
accountable ought to flow from an all things considered judgment factoring in considerations independent of the culpability of wrongdoers.

4.2 – What is restorative justice?

Basically, a restorative justice approach focuses on healing or repairing the harm of criminal wrongdoing, rather than punishing criminals for their misbehaviour (either because they deserve it, or for the sake of pursuing certain forward-looking aims, e.g. crime prevention, etc.). Partly in response to the deficiencies associated with more traditional approaches, defenders of restorative justice alternatives argue that the traditional approaches (e.g. consequentialism and retributivism) ignore or sideline important moral values. Advocates of restorative justice argue that by focusing exclusively on the justification for punishing wrongdoers (either to take retribution, or for the sake of pursuing certain forward-looking aims, e.g. crime prevention, etc.), the traditional accounts neglect an obvious reparative moral response to wrongdoing that has deep intuitive appeal, namely, the process of fostering moral recognition in wrongdoers toward the further aims of apology and reconciliation between

---

120 See Chapters 2 and 3. Also, there is some debate in the literature concerning the notion that restorative justice represents a response to more “traditional” approaches. Restorative justice theorists have suggested that restorative justice has deep historical roots. On this point, see Consedine, J., *Restorative Justice: Healing the Effects of Crime* (New Zealand: Ploughshares Publications, 1995). For a critical discussion of the point that restorative justice has deep historical roots, see Daly, “Restorative Justice: The Real Story,” pp.61-4. In my view, since I will be focusing on the merits of a restorative justice approach from a normative perspective, it is for my purposes irrelevant whether such an approach is historically newer or older than (what I describe in what follows as) the “traditional” approaches, e.g. consequentialism, retributivism, etc.
wrongdoers, their victims, and affected members of the community.  

Restorative justice is about undoing the damage done to the community resulting from crime.

The meaning of “restorative justice” is philosophically controversial. As mentioned in the introduction, two opposing views dominate the restorative justice literature. On the one hand, defenders of restorative justice argue that restorative justice stands in stark opposition to retributive justice, and therefore stands at odds with traditional approaches that, as opposed to moral reparation, focus exclusively on justifying a state-sponsored system of criminal punishment.  

Such theorists argue that wrongdoing between individuals should be resolved between individuals, and that the involvement of the state is antithetical to the pursuit of restorative justice. On the other hand, retributivists have responded to the former view by arguing that criminal punishment is not necessarily at odds with the ends of restorative justice. On the contrary, defenders of restorative retribution have argued that retributive punishment is a necessary condition for achieving the moral reparation that is at issue in the restorative justice literature.  

A common theme among such theorists is that the moral community, with the state as its representative, has an obligation to censure criminal wrongdoers via a public institution of criminal punishment.

---

122 Ibid.
123 For the earliest example of this sort of view, see Hegel, G.W.F., The Philosophy of Right, trans. T.M. Knox (Oxford: Oxford University Press, 1942).
124 See Duff, Punishment. Also see Duff, “Restorative Punishment.”
In the next section (section 3), I explain the merits of the former view, which I refer to in what follows as the *informal model of restorative justice*. I argue that this account points to significant weaknesses in many presently existing systems of criminal justice. Subsequently (section 4), I explain the retributivist response to the informal model, highlighting the argument that the community has an obligation to hold criminal wrongdoers accountable for the commission of public wrongs. Finally (section 5), I argue that retributive restoration presupposes a particular retributivist account of moral responsibility and blame. In response to this view, by appealing to a non-retributivist theory of moral responsibility and blame, I argue that we should develop a middle ground position on the relationship between restorative justice and criminal punishment. In my view, restorative justice is neither necessarily at odds with criminal punishment, nor does it necessarily require criminal punishment. Restorative justice does require, however, holding wrongdoers accountable to the community for their wrongful acts.

### 4.3 – The informal model of restorative justice explored

Defenders of restorative justice argue that the traditional approaches ignore or sideline important moral values. In order to appreciate the potential

---

125 Others have suggested that criminal punishment, although not necessary for restorative justice, can be reconciled with restorative values. See, for example, Strang, H., “Repair or Revenge,” in *A Restorative Justice Reader 2nd edition* ed. Gerry Johnstone (New York: Routledge, 2013) p.347. Strang writes, “Arguably, punishment as an outcome is not irreconcilable with restorative values…”
merits of a restorative justice alternative, it is worth entertaining a somewhat lengthy example. Consider the following case:\(^{126}\):

**Drunk Driver:** After a hard day at the office, Brian’s colleagues convince him to join them for a drink. Somewhat reluctantly, Brian agrees, although he is aware that he will need to drive home afterwards. After a short while, Brian gets up to leave, after having consumed the maximum amount of alcohol given that he needs to drive home. But the persuasion of his friends changes his mind, and he decides to stay for one more. Brian is not a reckless person by nature. He simply uses bad judgment in this instance, having had one drink too many before getting in his car to drive home. While he is not extremely intoxicated, he realizes he’s had too much to drink. Nevertheless, Brian recklessly proceeds to get in his car and begin making his way home. On the way, a cyclist named Jody is happening along as Brian is turning through an intersection. Jody, a somewhat inexperienced cyclist, falters at a critical moment, leaving her vulnerable in the middle of the road. Under normal conditions, Brian, who is turning through the intersection at the critical moment, should have been able to stop his vehicle or avoid her, but his reactions are delayed due to his slightly inebriated state. He attempts to swerve to avoid hitting Jody, but he fails, and his car makes direct contact with her, knocking her from her bike. Alcohol aside, it is uncontroversial that Brian is the one who is at fault in the accident. Luckily for both parties, Jody survives the accident. However, she is seriously injured. She sustains two broken legs and many more minor injuries, mostly scratches and bruises. Brian is immediately horrified by what he has done. He gets out of his car and approaches Jody, who is sprawled on the street and in severe pain. Brian’s immediate reaction is to apologize to Jody, although he also immediately realizes his apology is grossly inadequate given the situation. He also immediately wishes he could do more for Jody in order to make things right. Realizing the futility of his immediate inclinations, Brian promptly calls 911. Moments later, Brian finds himself in the back of a police car on his way to the police station, while at the same time, an ambulance takes Jody away to the hospital.

Let us assume that Brian has broken the law against impaired driving. Also, most would agree that Brian’s decision to drive while under the influence of alcohol is morally wrong. He is guilty of recklessly driving under the influence of alcohol, and by means of his recklessness and drunk driving, seriously injuring Jody, an

---

\(^{126}\) I borrow this hypothetical case from the example of Bryson and Judith in the introduction to Bennett, *The Apology Ritual*, pp.1-2.
innocent victim.\textsuperscript{127} What should be the response to Brian’s wrongful conduct? Such a question, we might think, involves two aspects: on the one hand, we can ask what should be Brian’s response to his own wrongdoing? On the other hand, we can ask what should be the response on the part of the community to Brian’s wrongdoing?

Traditional approaches to the problem of punishment focus almost exclusively on the latter question. Consequentialists focus on the forward-looking benefits of inflicting punishment upon Brian, either as a means to preventing crime, to rehabilitating or reforming Brian, to furthering the aim of social cohesion, etc. Retributivists, by contrast, are interested in what Brian deserves to suffer for his wrongful conduct. Retributivists focus on what the community has an obligation to do to Brian as the intrinsically appropriate response to his crime, e.g. to subject him to a proportional punishment. More recently, however, a newer approach to the problem of criminal wrongdoing focuses instead, at least primarily, on the appropriate response to wrongdoing on the part of the wrongdoer.\textsuperscript{128} Brian immediately recognizes that he is guilty of wronging Jody, and his immediate response involves an attempt to convey to her an apology (despite his realization that a merely verbal apology is going to be insufficient). Nevertheless, it is plausible that Brian’s own response to his wrongdoing in this case is critical to the moral architecture of the scenario. It is because he

\textsuperscript{127} It is possible to distinguish between the moral wrong on the part of Brian associated with driving under the influence of alcohol as such, and the moral wrong on the part of Brian associated with harming Jody by means of his recklessness and drunk driving. Cf. Kumar, R., “Who Can Be Wronged?” Philosophy and Public Affairs 31:2 (2003) p.103.
\textsuperscript{128} See Bennett, The Apology Ritual. Also see above, note 117.
immediately recognizes that he has wronged Jody, and immediately feels compelled to apologize to her, that we can consider Brian, despite his wrongdoing, a decent human being.\textsuperscript{129}

A restorative justice approach to the problem of criminal wrongdoing focuses primarily on repairing the damage done by criminal forms of wrongdoing. A natural point of departure, therefore, is to focus on the familiar moral practice of apology, i.e. the act/process of making amends via penitence of some sort. A critical claim often made by defenders of restorative justice is that more traditional approaches involve blocking or preventing the process of apology and moral reparation that might occur naturally without intervention, were it not for the involvement of the state, and the nature of the bureaucracy surrounding presently existing systems of criminal justice. For example, presently existing systems of criminal justice often sever all contact between wrongdoers and their victims. Such systems instead focus on legalities, procedures, and adversarial modes of conflict resolution. Once Brian is taken away by the police, he will be introduced to a system that, as Bennett argues, involves “…its own procedures, assumptions and language: in short its own culture or way of doing things.”\textsuperscript{130}

Brian will be charged, perhaps tried in a court of law, and subjected to punishment as the law prescribes. He will have the opportunity, if he decides to pursue it, to challenge the charges against him, and if he is convicted, to appeal

\textsuperscript{129} Bennett makes a similar point about Bryson in his example that I am borrowing: “…we can all recognize the appropriateness of … [Bryson’s] response. It is because Bryson feels compelled to … [respond in this way] that we can regard him as, despite his misdemeanor, a basically decent human being.” See Bennett, \textit{The Apology Ritual}, p.3.

\textsuperscript{130} See Bennett, \textit{The Apology Ritual}, p.2.
his conviction. In short, he is entered into a fundamentally adversarial system that seems to pit his interests against the interests of others in the community, including his victim. Critics of traditional approaches to criminal justice have pointed out that if the charges against Brian carry excessive threats to his personal wellbeing (e.g. a lengthy term of confinement in prison), then he will be motivated to respond to the scenario with self-interest as opposed to the response of apology and moral atonement that is his initial, natural reaction.131

However, another approach to responding to the incident involving Brian and Jody is possible, an approach that I will refer to in what follows as the informal model of restorative justice: defenders of restorative justice often suggest that as a community, our responses to wrongdoing ought to build upon rather than impede the natural response of the virtuous offender, such as Brian’s response to his own wrongdoing. Brian immediately feels sorry for what he has done, and he immediately has the impulse to apologize to Jody. He also has the impulse to do something (or many things) for her in order to imbue his apology with meaning, toward the further end of rectifying the situation to whatever extent possible, i.e. repairing the harm caused by his wrongdoing. Critics of traditional approaches to criminal justice often point out that it is this fundamentally restorative moral process that presently existing systems of criminal justice tend to inhibit.132 Therefore, defenders of restorative justice argue that our responses as a community to criminal forms of wrongdoing ought to build upon, rather than inhibit, the moral process of apology and reconciliation that Brian and other

131 For this point see Bennett, *The Apology Ritual* p.3.
virtuous offenders feel compelled to pursue of their own accord. Put in stronger terms, advocates of a restorative justice approach argue that restorative justice takes as its point of departure the responsibilities of wrongdoers to recognize what they have done as wrong, and to make amends.  

But what, one might ask, would such an alternative approach prescribe? For example, what should happen to Brian rather than being formally charged and punished by the state, as presently existing systems of criminal justice typically prescribe? In response to this question, defenders of restorative justice often argue that a process of “deliberative dialogue” should replace, to whatever extent possible, presently existing adversarial systems of state punishment. Braithwaite and Strang distinguish between two conceptions of restorative justice, the process conception and the values conception. Describing the process conception, they write,

On this view, restorative justice is a process that brings together all stakeholders affected by some harm that has been done (e.g., offenders, their families, victims and their families, affected communities, state agencies such as the police). These stakeholders meet in a circle to discuss how they have been affected by the harm and come to some agreement as to what should be done to right any wrongs suffered.

The process conception identified by Braithwaite and Strang therefore emphasizes a procedural departure from more traditional, state-sponsored

systems of criminal punishment. Rather than bringing offenders before the state to be tried and punished, offenders are viewed as a single “stakeholder” within a larger community affected by wrongdoing, including the victim(s), all of whom should have a part in determining “what should be done to right any wrongs suffered.” In the case of Brian, therefore, all affected members of the community (including Brian and Jody, their families, but also the police, first responders, witnesses of the incident, members of the community who live in the area, etc.) should come together, along with a mediator, to discuss their various perspectives on the incident and its implications (moral, material, psychological, etc.), and to come to an agreement regarding what should be done going forward to right any wrongs suffered. All parties might agree, for example, that Brian ought to give up his driver’s license for a lengthy period of time, give lectures at local schools in the community regarding dangerous driving and driving under the influence, and agree to aid in Jody’s recovery in certain ways, perhaps in part by fully compensating her for the material losses associated with her injuries.136

According to a values conception of restorative justice, Braithwaite and Strang write,

…it is values that distinguish restorative justice from traditional punitive state justice. Restorative justice is about healing (restoration) rather than hurting. Responding to the hurt of crime with the hurt of punishment is rejected, along with its corresponding value of proportionality – punishment that is proportionate to the wrong that has been done. The idea is that the value of healing is the key because the crucial dynamic to foster is healing that begets healing. The dynamic to avert is hurt that begets hurt.137

136 For a similar point see Bennett, The Apology Ritual, p.6.
137 See Braithwaite and Strang, Restorative Justice and Civil Society, pp.1-2.
Therefore, according to both conceptions of restorative justice outlined above, as well as a plethora of arguments made by other defenders of restorative justice alternatives,\(^\text{138}\) restorative justice stands in stark opposition to retributive justice, or “traditional punitive state justice.”\(^\text{139}\) Part of the objection to more traditional state-oriented approaches to criminal justice is that wrongdoers should be held accountable to their victims and other affected members of the community, rather than to the state. According to presently existing systems, Brian is made to “pay” for his crime (e.g. via a hefty fine, imprisonment, community service, probation, etc.) in ways that seem directed towards the state rather than towards Jody.\(^\text{140}\)

However, rather than bringing wrongdoers before the state, the informal model of restorative justice aims to address wrongdoing to whatever extent possible as a private matter, to be resolved between wrongdoers, their victims, and affected members of the community.\(^\text{141}\) The informal model aims to bring together all the stakeholders involved in the crime in order to discuss their

---


\(^{139}\) See Braithwaite and Strang, *Restorative Justice and Civil Society*, pp.1-2. Also see the introduction to Bennett, *The Apology Ritual*.

\(^{140}\) See Bennett, *The Apology Ritual*, pp.3-4.

\(^{141}\) As will be discussed in what follows, it is a matter of ongoing debate whether or not restorative justice processes ought to be considered private as opposed to public, in the sense of attempting to exclude state involvement in the pursuit of restorative justice to whatever extent possible. For the claim that (what I refer to as) the informal model of restorative justice aims at pursuing criminal justice as to whatever extent possible a private rather than public matter, see Bennett’s discussion of the “laissez-faire conception of restorative justice,” Bennett, *The Apology Ritual*, pp.126-133. Cf. Zehr, H. *Changing Lenses*. 
differing perspectives, to foster recognition and understanding of the implications of the wrongdoing in question, and to determine what ought to be done to “right the wrong” to whatever extent possible. Rather than focusing on the backward-looking aim of retribution, therefore, the informal model focuses primarily on certain forward-looking aims, such as benefiting the victim(s) and affected members of the community, restoring and reintegrating the wrongdoer, and repairing the damage done by the wrongdoing in question to whatever extent possible. Along these lines, the informal model allows for wrongdoers, victims, and affected members of the community to exercise a degree of control over the criminal justice process.  

Furthermore, the informal model emphasizes moral recognition, in the sense of allowing the stakeholders involved to have their voices heard, and to share their differing perspectives on what has happened, and what ought to be done going forward. Finally, the informal model attempts to foster understanding, in that by allowing for dialogue between the wrongdoer, the victim(s), and affected members of the community, it is more likely that all affected parties will come to a better understanding of the implications (moral, material, psychological, etc.) of the instance of wrongdoing in question for others, and how best to proceed going forward in order to resolve the situation to whatever extent possible.

Given the deficiencies associated with the traditional approaches, it is possible to situate in context the (at least prima facie) appeal of the informal

---

model of restorative justice. Recall that, in the pursuit of certain valuable forward-looking aims, traditional consequentialist approaches fail to treat wrongdoers with appropriate respect. Traditional consequentialist approaches are open to the objection that they aim for valuable outcomes by means of a system that involves the use of mere force and control. The informal model of restorative justice, by contrast, pursues valuable forward-looking aims, including fostering moral recognition in wrongdoers, and reconciliation between wrongdoers and their communities, but attempts to do so in a way that treats wrongdoers as fully rational and responsible moral agents. The informal model aims at fostering recognition in wrongdoers via a deliberative process of moral communication and dialogue, rather than via a process of mere manipulation and control. We can therefore describe the informal model of restorative justice as prescribing a process of shared moral inquiry that aims at repairing the damaged relationships in the community resulting from crime.\footnote{See Bennett, \textit{The Apology Ritual}, pp.75-82.}

Also, recall that traditional versions of retributivism arguably fail to adequately explain why hard treatment is necessary in all cases in order to hold wrongdoers accountable to others. The informal model of restorative justice, by contrast, rejects this backward-looking aspect of retributivism, i.e. that hard treatment is something that wrongdoers, as a matter of justice, deserve to suffer. However, the informal model nevertheless attempts to hold onto the notion that wrongdoers ought to be held accountable to their victims and others. Also, as mentioned above, the informal model aims to treat wrongdoers with the respect
they are owed as rational and responsible moral agents, i.e. as moral agents that should be involved in a process of shared moral inquiry. If a form of hard treatment (e.g., a period of confinement, community service, etc.) is the agreed upon output of a deliberative process that brings together all stakeholders in the crime, then it is not imposed on the wrongdoer as something she deserves to suffer per se, but rather as an aspect of the forward-looking aims of fostering recognition and making amends, and thereby restoring the relationships between wrongdoers and their victim(s), and others in the community negatively affected by the crime(s).

4.4 – A retributivist response to the informal model of restorative justice

The informal model of restorative justice has inspired some retributivist thinkers to defend a version of “restorative retribution,” or “retributive restoration.” In response to the informal model, retributivists argue that criminal forms of wrongdoing constitute “public wrongs,” rather than merely “harms” that negatively affect specific members of the community. In other words, criminal forms of wrongdoing are not only the appropriate business of specific, directly affected members of the community. Criminal forms of wrongdoing are also the appropriate business of the community as a whole. Retributivists have argued that domestic violence is a relevant example here. Domestic violence is an example of a “public wrong,” in the sense that the condemnation of domestic violence

---

violence is appropriately the business of the community as a whole. To fail as a community to condemn domestic violence would be a further, distinct form of collective wrongdoing.\footnote{For this point see Bennett, \textit{The Apology Ritual}, p.142.}

In response to the informal model, therefore, retributivists have argued that, due to the nature of criminal forms of wrongdoing, we should view a state-sponsored system of retributive punishment as a necessary condition for the pursuit of restorative justice. Kathleen Daly writes, “When one first dips into the restorative justice literature, the first thing one ‘learns’ is that restorative justice differs sharply from retributive justice.”\footnote{Daly, “Restorative Justice: The Real Story,” p.58.} However, according to Daly, we should reject the stark distinction between restorative justice and criminal punishment presupposed by the informal model:

\begin{quote}
I have come to see that apparently contrary principles of retribution and reparation should be viewed as dependent on one another. Retributive censure should ideally occur before reparative gestures (or a victim’s interest or movement to negotiate these) are possible in an ethical or psychological sense.\footnote{Daly, “Restorative Justice: The Real Story,” p.60.}
\end{quote}

Along similar lines, Duff argues,

\begin{quote}
…once we gain a better understanding of the concepts of restoration and of punishment, we will be able to dissolve the apparent conflict between them, and to see that criminal punishment should aim at restoration, whilst restorative justice programmes should aim to impose appropriate kinds of punishment.\footnote{Duff, “Restorative Punishment,” p.368.}
\end{quote}

Therefore, for both Daly and Duff, a state-sponsored system of criminal punishment is not necessarily at odds with restorative justice. Rather, both Daly
and Duff argue that moral reparation, in the sense that is at issue in the restorative justice literature, requires retributive punishment.

But, one might ask, what makes retributive punishment, as Duff suggests, necessary to pursuing the ends of restorative justice? Answering this question requires that we delve deeper into Duff’s view concerning how to understand the concepts “retribution” and “restoration.”

According to Duff, generally speaking, many things are potential candidates for restoration, including a person’s property, health, or reputation. In certain cases, the aim of restoration is to achieve what Duff describes as the “status quo ante,” or the state prior to some harm inflicted, such as damage or illness. Of course, it is not always possible to fully restore, for example, someone’s property or health. Property might be damaged or altered beyond repair, or have intrinsic value that goes beyond its immediate material value, e.g. a family heirloom. Likewise, certain injuries to the body may never fully heal, such as certain permanent forms of physical or psychological injury. In such cases, while full restoration is impossible, instead, partial reparation might take the form of compensation intended to offset irreparable damage. Reparation in the form of compensation is made necessary when it is impossible to restore fully whatever has been lost through injury.

However, damage done to one’s property or health may or may not be the result of culpable wrongdoing. Therefore, we must inquire as to the appropriate

\[153\] Ibid.
\[154\] Ibid.
metric for restoration in cases of wrongdoing as opposed to mere harm, and in particular, in cases of criminal forms of wrongdoing. In the case of criminal wrongdoing, Duff argues, rather than aiming for some form of material or psychological restoration, the ends of restorative justice involve restoring the normative relationship between wrongdoers and their communities (and between wrongdoers and victims as co-members of their communities).\textsuperscript{155} He writes, “The damage to the relationship must … be understood in normative, not merely empirical, terms.”\textsuperscript{156} In this sense, Duff argues we should be willing to scrutinize, for example, victims of wrongdoing who are in fact willing to accept an apology that is insufficiently sincere, or inappropriately weighty given some set of circumstances, e.g. given an egregious form of wrongdoing. Conversely, we should be willing to scrutinize those who are in fact unwilling to accept any form of apology or reparative gesture, given some set of circumstances, e.g. in a case involving a relatively minor offence. What is at issue, therefore, is what will repair the normative relationship, i.e. the relationship between the wrongdoer and the victim as co-members of a shared normative community, rather than, for example, the psychological fact of whether or not some particular victim happens to accept (or refuse to accept) an apology from some particular wrongdoer.\textsuperscript{157}

On Duff’s view, we should see communicative punishment – i.e. punishment that aims at bringing about penitential reform in the wrongdoer – as

\textsuperscript{155} See Duff, “Restorative Punishment,” p.371.
\textsuperscript{156} Ibid.
\textsuperscript{157} Ibid.
necessary to repairing the normative relationship between the wrongdoer and her community in at least three distinct ways:

[T]he kind of mediation appropriate to crime … is indeed a retributive process: for if we ask why it is appropriate that she [i.e. the wrongdoer] should be thus burdened, why she should be brought to suffer, the answer is … that this is what she deserves for her crime. She deserves to suffer the censure of others, and her own remorse: for that censure and her remorse are appropriate responses to her crime. She [also] deserves to suffer the burden of making moral reparation for what she has done…¹⁵⁸

For Duff, therefore, retribution is an aspect of what is necessary to repairing the damaged relationships between wrongdoers and their communities in at least the following three senses: (1) criminal punishment, understood as a mode of forceful, public censure, metes out the blame or censure that wrongdoers deserve to suffer, as the intrinsically appropriate response on the part of the community to criminal forms of wrongdoing. Furthermore, (2) wrongdoers deserve to feel the pain of remorse, which is an aspect of the appropriate response on the part of the wrongdoer to her own wrongdoing. Finally, (3) the wrongdoer deserves to suffer the burden associated with her obligation to meaningfully apologize to her victims and the community, toward the further end of reconciliation between her and her victim(s) as co-members of the normative community. According to Duff, therefore, there are at least three distinct senses in which retributive punishment – and therefore also the deliberative processes emphasized by restorative justice theorists – should aim to “bring … [wrongdoers] to suffer what they deserve.”¹⁵⁹

¹⁵⁹ Ibid.
In this way, Duff’s view represents a middle-ground position between two opposing theories, namely, between (1) traditional “just deserts” retributivism, according to which there is no value in punishing wrongdoers beyond meting out the suffering they deserve, and (2) the informal model of restorative justice, according to which there is no value in state-sponsored criminal punishment whatsoever, since an institution of criminal punishment stands in stark opposition to the ends of restorative justice, the latter of which ought to be pursued over and above criminal punishment. For Duff, we should reject both views, and instead pursue the ends of restorative justice through retributive punishment, as a necessary condition for restoring the normative relationship between the wrongdoer and her victim(s) as co-members of a shared normative community.

But, one might ask, to what extent is retributive punishment really at issue in Duff’s formulation of what restorative justice requires? Return for a moment to Brian and Jody in the Drunk Driver case. It is true, we might think, that Brian has an obligation to feel remorse, in the sense of recognizing that what he has done to Jody (and others in the community) was wrong. Furthermore, it seems plausible to hold that Brian owes it to Jody and others to make amends, in the sense of shouldering certain burdens (e.g. giving up his drivers license, giving lectures in the community, materially compensating Jody, etc.) in order to demonstrate his remorse, and pursue a process of restorative reconciliation between himself and those whom he has wronged. It seems plausible, therefore, that the aims of fostering recognition and reconciliation can be accommodated by the informal model that eschews retributive punishment. Furthermore, it seems
conceptually implausible to construe the burdens of moral recognition and apology after the fact of wrongdoing as burdens the wrongdoer deserves to suffer. It seems more conceptually plausible to construe the burdens of moral recognition and apology in terms of the duties of the wrongdoer.\textsuperscript{160}

However, as outlined above, Duff and other retributivists also claim that wrongdoers deserve to suffer the condemnation of the moral community as a whole. This is, I submit, the most straightforward sense in which wrongdoers “deserve to suffer” according to the communicative theory of punishment; wrongdoers deserve to suffer public censure, as the appropriate response to criminal forms of wrongdoing on the part of the normative community. To sharpen the intuition, we might imagine another case – call it Drunk Driver 2 – wherein Alfred, a thoroughly unrepentant offender, is in the same situation as Brian in Drunk Driver, however Alfred is stubbornly unwilling to recognize that what he has done is wrong, let alone willing to voluntarily make amends for his crime. In such a case, we might think that Alfred deserves to suffer an expressive or communicative form of criminal punishment that aims at censuring him and condemning his wrongful conduct on behalf of the victim, those directly affected by his conduct, and the wider, normative community.

\textsuperscript{160} The construal of the burdens of moral recognition and apology as burdens the wrongdoer deserves to suffer misleadingly characterizes such burdens as the sorts of things that can be imposed upon the wrongdoer independently of her agency. Understood as duties, by contrast, it is apparent that, from the perspective of the wrongdoer, recognizing wrongdoing as wrongdoing and shouldering certain burdens for the sake of making amends are the sorts of things wrongdoers, by virtue of wronging others, are obliged to take it upon themselves to do. This point will be taken up in Chapter 5.
In my view, therefore, retributivists are correct that the informal model of restorative justice neglects an important collective obligation on the part of the normative community as a whole, i.e. the obligation to hold wrongdoers accountable for the commission of public wrongs. Must we accept, then, that restorative justice requires retributive punishment, in the sense described above, i.e. that wrongdoers deserve to suffer expressive or communicative punishment that aims at censuring their wrongful conduct? I return to this question in the following section. It is true that for Duff and others, communicative punishment is constitutive of the ends of restorative justice. Communicative punishment – understood as a “penitential process” imposed upon the wrongdoer – is not only necessary for the wrongdoer to repair her relationship with the normative community; for Duff, communicative punishment constitutes the sort of moral reparation that is at issue between wrongdoers and their communities. In the following section, my aim is to sketch out an alternative approach. In my view, holding wrongdoers in an appropriate way accountable is the appropriate response to criminal forms of wrongdoing on the part of the normative community, which may or may not take the form of communicative punishment. Famously, censuring wrongdoers and their wrongful conduct does not necessarily involve the imposition of hard treatment.\textsuperscript{161} Purely formal or symbolic modes of censure might take the place of hard treatment, thereby satisfying the condition that the wider, normative community has an obligation to censure criminal wrongdoers. By appealing to a non-retributivist account of moral

\textsuperscript{161} For this point see Feinberg, “The Expressive Function of Punishment” p.423.
responsibility and blame, my aim in the following section is to sketch out an alternative non-retributivist account of moral responsibility and restorative justice, that explains why holding wrongdoers accountable to the community, at least in certain cases, does not require that the wrongdoer undergo a process of retributive punishment per se.

4.5 – Moral responsibility, blame, and restorative justice

In my view, restorative retributivists are correct to criticize the informal model of restorative justice for failing to take seriously the claim that the community has an obligation to hold wrongdoers accountable, in the sense of censuring wrongdoers for the commission of public wrongs. Criminal forms of wrongdoing are, by their very nature, the appropriate business of the community as a whole. Therefore, the informal model falls short in at least one important respect: the prescribed process of “deliberative dialogue,” insofar as it is construed as standing in opposition to a state-sponsored response to criminal forms of wrongdoing, sidelines the important consideration that those members of the community not directly affected by the wrongdoing in question nevertheless have a stake in its condemnation. 162 Domestic violence, for

---

162 See above, note 141. As discussed above, the view that restorative justice stands in stark opposition to a state-sponsored system of criminal punishment is a widely held view in the restorative justice literature. In what follows, I argue that the anti-statist dimension of restorative justice presupposed by the informal model ought to be rejected. While the informal model of restorative justice would be adequate in dealing with certain forms of wrongdoing in a thoroughly private setting (i.e. in certain organizational contexts, e.g. wrongdoing between colleagues within an academic department), insofar as an instance of wrongdoing between individuals constitutes a criminal form of wrongdoing, i.e.
example, seemingly private though it may be in its character as an instance of wrongdoing, is something that all members of the community should have a voice in condemning.\footnote{163}

However, in my view, we should question the further claim made by restorative retributivists that criminal punishment is the only way for the community to hold wrongdoers accountable and censure their wrongdoing. As Feinberg famously argues, public censure might take a merely formal or symbolic form, such as a public admonishment of some sort.\footnote{164} The criminal conviction itself is an example of merely formal, public censure. Duff’s view seems to presuppose a strongly retributivist theory of moral responsibility and blame. Not only does the community have an obligation to censure criminal wrongdoers; the community has an obligation to mete out communicative punishment – understood as hard treatment that aims at bringing about penitential reform in the wrongdoer – as the intrinsically appropriate \textit{mode} of censuring wrongdoers for constitutes a public wrong as determined by the political community (how the scope of the criminal law ought to be determined will be further discussed in what follows), in my view, responding to crime with some form of condemnatory response falls under the appropriate purview of the state. In this respect, I am in agreement with Duff and Bennett. See Duff, \textit{Punishment} and Bennett, \textit{The Apology Ritual}. As I argue in what follows, however, where I disagree with the restorative retributivists is in the context of the claim that criminal punishment is in all cases the necessary condemnatory response to crime on the part of the community as a whole. While criminal wrongs by their very nature require a condemnatory response of some sort on the part of the community as a whole, there are reasons to suppose that criminal punishment is not always the most appropriate condemnatory response.

\footnote{163} See Duff, \textit{Answering for Crime}, pp.140-6. Also see Bennett, \textit{The Apology Ritual}, p.142.

\footnote{164} Feinberg, “The Expressive Function of Punishment” p.423.
their criminal acts.\textsuperscript{165} It is therefore worth examining in greater detail the question of whether or not restorative justice requires a retributivist account of moral responsibility and blame.

First, consider Duff’s view: Duff’s commitment to retributivism as a constitutive element of restorative justice hinges upon his commitment to an essentially retributivist view of moral blame and its relationship to restorative justice. For Duff, forceful public censure, \textit{in the form of communicative punishment}, is the appropriate backward-looking response to criminal forms of wrongdoing on the part of the community.\textsuperscript{166} Communicative punishment is necessary to restore the normative relationship between the wrongdoer and her victim as co-members of the normative community.\textsuperscript{167} Therefore, in response to Duff, if we are unsure whether or not restorative justice requires retributive punishment, we should look to a contemporary non-retributivist account of moral responsibility and blame, in order to see how such an account holds up in the context of criminal responsibility and restorative justice.

Most basically, we can ask, what does it mean to “hold someone responsible” for their wrongful conduct? As we have seen, for Duff, communicative punishment is necessary in order to appropriately “hold wrongdoers responsible” to their communities, in a way that is required for

\textsuperscript{165} Cf. von Hirsch, \textit{Censure and Sanctions}. For von Hirsch, the criminal sanction involves public censure plus a prudential disincentive in the form of penal hard treatment. See Chapter 3 section 4. For Duff, by contrast, the hard treatment involved in the criminal sanction is constitutive of the censure that wrongdoers deserve.

\textsuperscript{166} See Duff, \textit{Punishment}, pp.81-2.

\textsuperscript{167} See Duff, “Restorative Punishment,” p.380.
restorative justice. However, in making sense of this aspect of Duff’s account, as well as potential alternatives to Duff’s account, we must unpack the idea of moral blame and its relationship to “holding wrongdoers responsible.” Importantly, retributivism about moral responsibility and blame is not the only available view concerning the normative significance of “holding wrongdoers responsible.” On Scanlon’s view, for example,

[T]o claim that a person is blameworthy for an action is to claim that that action shows something about the agent’s attitudes toward others that impairs the relations that others can have with him or her. To blame a person is to judge him or her to be blameworthy and to take your relationship with him or her to be modified in a way that this judgement of impaired relations holds to be appropriate.\(^{168}\)

Now, since Scanlon’s comments about blame in this context concern blame as it relates to inter-personal relationships (as opposed to impersonal relationships, as in the context of our discussion), we should tread carefully, so as not to misconstrue the relevance of Scanlon’s position to our discussion in what follows. Nevertheless, Scanlon’s account of moral responsibility and blame suggests an attractive alternative to retributivism about moral responsibility and blame.

At this juncture, in order to set out an alternative to Duff’s account of moral responsibility and blame as it relates to restorative justice, it is necessary to review some basic features of Scanlon’s view. First, for Scanlon, moral responsibility and the appropriateness of blame ought to be considered grounded in the relationships we have with others. These relationships (e.g. as friends, family members, co-citizens, etc.) ground the normative standards for how we ought to relate to each other, for what sorts of attitudes and actions ought to be

considered appropriate or permissible in the context of such relationships, for what we owe to each other, etc. Second, when such standards are violated through wrongdoing, the relationships between individuals (i.e. between the wrongdoer and the one who is wronged, but also between the wrongdoer and others) suffer *impairment*. Third, for Scanlon, the *standing* of the blamer in relation to the one who is blamed is relevant to the significance or meaning of the impairment, as well as the appropriate response to the impairment on the part of the blamer.

For example, suppose that Frances and Gertrude are friends. As friends, we can assume that Frances and Gertrude should support each other in their personal projects.¹⁶⁹ Now suppose that Frances consistently fails to support Gertrude in her personal projects. Perhaps Frances often ridicules Gertrude for pursuing certain personal projects, and undermines her efforts to achieve her personal goals. Frances’s attitudes and actions in this context impair their friendship. However, the meaning of the impairment and the appropriate response on the part of Gertrude will depend upon Gertrude’s standing in the relationship. If, for example, Gertrude is also guilty of undermining Frances’s personal projects in a similar way, then this fact will affect the meaning of the impairment for Gertrude and others, and the appropriate response on the part of Gertrude to the impairment. Even if their friendship is fundamentally damaged beyond repair as a result of the respective undermining attitudes and actions of

¹⁶⁹ I am assuming, here, that supporting one’s friends in their personal projects is a plausible normative standard grounded in the relationship of being someone’s “friend.”
both parties, it would seem in this case hypocritical for Gertrude to respond to
Frances’s behaviour with, for example, righteous indignation (as she might
reasonably respond if the circumstances were otherwise, i.e. if Gertrude had not
also violated the standards of friendship in a similar fashion). Perhaps a more
appropriate response on the part of Gertrude in this case would be to recognize
her own failures associated with the friendship, and to withdraw (in permissible
ways) from the relationship. Notice, however, that it would not be incorrect for
Gertrude to judge Frances’s behaviour as blameworthy, and to blame her for
failing to live up to the normative standards appropriate to friendship. However,
Gertrude’s standing in the relationship does affect how she ought to actively
respond, i.e. the expression of blame appropriate to the context.

Along these lines, in the context of criminal responsibility and blame, what
I want to focus on in what follows is the distinction between judging a wrongdoer
to be blameworthy, on the one hand, and the expression of such a judgement by
various possible means, on the other. As Smith argues, we can break down the
idea of “holding wrongdoers responsible” into at least three distinct aspects
\(^\text{170}\):

(1) Those who have been wronged (X) hold the wrongdoer (Y) to be the agent
responsible for the act or attitude in question.

(2) X holds Y to be culpably or wrongfully responsible for the act or attitude in
question.

\(^\text{170}\) Smith, A., “On Being Responsible and Holding Responsible,” The Journal of
(3) And finally, X actively blames Y for the act or attitude in question.

According to (1), holding wrongdoers responsible requires that the agent is, in fact, the agent to which we can attribute the act or attitude in question. It is plausible that the fact of being responsible for some act is a condition for being held responsible for it by others. Concerning (2), beyond merely attributing to the agent the act / attitude in question, the wrongdoer must be deemed culpable, or wrongfully responsible for the act / attitude in question. It is plausible that both (1) and (2), i.e., both being judged to be responsible and being judged to be culpably responsible, are necessary conditions for the permissibility of being actively blamed by others for some act. With respect to (3), holding wrongdoers responsible might be understood as an expression of the judgements of responsibility and culpable responsibility in (1) and (2). Therefore, (3) entails that (1) and (2) have been satisfied. However, as Smith argues, (3) is not entailed by (1) and (2). It is possible to judge a wrongdoer to be culpably responsible for some wrong act without actively blaming her for such an act.

Furthermore, as Smith argues,

The judgement that a person is morally culpable for some [act or] attitude … forms an essential part of the justification for … blaming attitudes and responses. But … the question whether it would be appropriate for any particular person to have and express any particular attitude to the agent will depend upon many considerations in addition to the agent’s responsibility and culpability for the thing in question.\(^\text{171}\)

Smith argues that the reasons why we might withhold or alter the expressive dimension of blame may have little or nothing to do with the judgement that the wrongdoer is indeed culpably responsible for some act. For example, the blamer might lack the appropriate standing in relationship to the wrongdoer, or alternatively, the blamer might take into account the wrongdoer’s response to her own wrongdoing as a mitigating or overriding factor.

Likewise, in the context of criminal responsibility and blame, I submit that we should distinguish between judgements of culpable responsibility and blameworthiness, on the one hand, and the appropriate expression of such judgements in whatever form, on the other. As Smith suggests, in my view, the appropriate mode of expressing the judgement that some wrongdoer is culpably responsible for some act / attitude ought to be determined in part by appeal to considerations that are independent of the culpable responsibility of the offender as such.

As we have seen, for Duff, the intrinsically appropriate response to criminal forms of wrongdoing on the part of the normative community is criminal punishment that communicates the censure that the wrongdoer, by virtue of being a wrongdoer, deserves to suffer. He argues that such a response is necessary to the pursuit of restorative justice, understood as repairing the normative relationship between the wrongdoer and her community. However,

---

174 I am assuming, here, that in the context of criminal responsibility, some form of expression of the judgement that the wrongdoer is blameworthy is appropriate, thereby satisfying the condition that the community has an obligation to hold wrongdoers accountable for the commission of public wrongs.
while I agree with Duff that we should consider giving expression in some way to the judgment that the wrongdoer is culpably responsible for wronging others as necessary to pursuing the ends of restorative justice, we should question his more specific claim that communicative punishment is the necessary mode of giving expression to such a judgement in all cases, for the sake of pursuing such ends. We should not only distinguish, as Smith does, between judgements of blameworthiness and the expression of such judgements, we should also distinguish between the various possible modes of expressing judgements of blameworthiness (i.e. modes of holding wrongdoers accountable, and thereby expressing condemnation) at the disposal of the normative community, and appropriate to various contexts.

Now, importantly, this is not to claim that we should let wrongdoers “off the hook,” so to speak, for their criminal acts. In my view, as restorative retributivists rightly suggest, wrongdoers, by virtue of wronging others, have obligations to recognize their wrongdoing for what it is, namely, wrongdoing, and to attempt to make amends by various available means. The more modest claim I am advancing is that communicative punishment, as Duff understands it, i.e. as hard treatment that aims at bringing about penitential reform on the part of the wrongdoer, is not the only possible mode of expressing the judgement that the wrongdoer is blameworthy and hence ought to be held accountable for her

---

175 The wrongdoer’s duties of moral recognition (in both the attitudinal and expressive senses outlined here) will be taken up further in the following Chapter (Chapter 5), and make up the basis for my argument for why it is permissible (although not obligatory) to impose hard treatment on wrongdoers for the sake of restorative justice.
crimes to the wider, normative community. We might judge it to be appropriate in some case that merely formal censure, e.g., in the form of a criminal conviction (but without further penalty), is the preferable mode of expression, all things considered. Or, alternatively, we might judge it to be appropriate in some case that merely formal censure (e.g. in the form of a criminal conviction) ought to be coupled with a process of deliberative dialogue of the form prescribed by the informal model of restorative justice. Such alternatives would still meet the demand that wrongdoers ought to be held accountable to the wider, normative community for the commission of a public wrong.

As Smith suggests, our judgement concerning the appropriate mode of expressing blame ought to flow in part from considerations that are independent of the culpability of the wrongdoer. But, one might ask, in the context of criminal responsibility, what are some of the relevant considerations that fall outside of the wrongdoer’s culpable responsibility for some act / attitude that ought to play a

---

176 One possibility is that the court’s ruling might include sentencing suggestions constraining the process of deliberative dialogue (i.e. a range of suggestions including upper and lower limits in terms of severity).

177 Bennett suggests that the reason public censure must take the form of proportionate hard treatment is that the mode of censure must be “symbolically adequate.” See Bennett, *The Apology Ritual*, p.149. Bennett writes, “…condemnation has to be expressed in symbolically adequate terms…” However, in my view, as I argue in what follows, the symbolic adequacy of the form of expressed censure must also factor in considerations independent of the culpability of the wrongdoer. However, if “symbolic adequacy” is defined in such a way that excludes considerations independent of the culpability of the wrongdoer (i.e. if symbolic adequacy *means* hard treatment proportionate to the seriousness of the offence in question), then the symbolic adequacy of the hard treatment chosen to express censure must be understood as one consideration among others that ought to factor into our all things considered judgement concerning how best to express the judgement that the criminal wrongdoer in question is blameworthy and hence ought to be held accountable to the community. This argument will be further developed in what follows.
role in determining the appropriate mode of expressing blame? In other words, what are some of the other salient considerations that ought to factor into our collective judgement regarding the most appropriate mode of expressing the judgement that some wrongdoer is blameworthy, and hence ought to be held accountable to the community for her criminal act(s)?

I submit that, in addition to the culpability of the criminal wrongdoer, we should also, as Smith argues, take into consideration our standing as a community in relation to the wrongdoer, e.g. whether she is a particularly marginalized or otherwise disadvantaged member of the community. In such cases, while holding the wrongdoer accountable for wronging others is nevertheless appropriate, subjecting her to criminal punishment might not (or at least not always) be the most appropriate mode of holding her accountable and expressing condemnation. Alternatives to criminal punishment, such as mandatory engagement in a process of shared moral inquiry, might be a more appropriate mode of expressing condemnation in the case of socially and politically marginalized offenders, all things considered.

---

178 There are numerous well-known examples in the context of criminal justice wherein marginalized groups are overrepresented in the penal system. In Canada, for example, the representation of Aboriginal offenders in prisons is disproportionately high, i.e. at a ratio of 10:1 when compared with non-Aboriginal offenders. However, Aboriginal people in Canada are also a segment of society subjected to systematic oppression at the hands of the political community, in the sense of being deeply socially and politically marginalized. Macdonald writes, “In Canada, the Indigenous incarceration rate is 10 times higher than the non-indigenous population – higher even than South Africa at the height of apartheid.” See, for example, Macdonald, N., “Justice is not blind,” Maclean’s (February 29, 2016).
In addition, retributivism about blame has the potential to take for granted the status of the blamer in relation to the blamed, i.e. that the individual/community is not herself/itself guilty of wronging the wrongdoer, hence compromising the status of the blamer as a moral judge. This idea is well established in the philosophy of moral responsibility and blame, yet has received comparatively little attention in the philosophy of punishment. Retributivists about punishment often argue that certain social and political background conditions must first be satisfied in order for retributive punishment to be morally legitimate.\textsuperscript{179} However, the emphasis is often placed on the relevance of social and political background conditions as mitigating factors related to the culpability of the offender, rather than as mitigating factors related to the standing of the community engaging in the collective expression of blame. In a way analogous to the case of Frances and Gertrude, the fact of marginalization might alter the expression of blame on the part of the community that is appropriate to any given context of wrongdoing perpetrated by a marginalized offender. The standing of the community in relation to the wrongdoer is relevant to how the community ought to respond. A restorative justice approach seems particularly appropriate in the context of the problem of marginalized offenders; both the wrongdoing perpetrated by the individual offender and the collective wrongdoing perpetrated

\textsuperscript{179} See, for example, Murphy, "Marxism and Retribution." Importantly, for the traditional Marxist, no wrongdoing occurs when pernicious social and political background conditions are responsible for alienating the wrongdoer from herself and her community. Rather, according to traditional Marxism, “wrongdoing” under such conditions is not wrongdoing at all, but rather a form of mere harm resulting from malady. See Sypnowich, C., \textit{The concept of socialist law} (Oxford: Oxford University Press, 1989) Ch.1. Hence, on the traditional Marxist view, under such conditions the legitimacy of retributive punishment is compromised.
by the political community as a whole damage the civic relationship between the offender and her community.\textsuperscript{180} A restorative response to wrongdoing seems particularly appropriate once we take seriously the notion that the political community as a whole may be guilty of perpetuating injustice.\textsuperscript{181} Indeed, the ends of restorative justice (understood in terms of the aims of fostering recognition in wrongdoers and reconciliation between wrongdoers and their communities) demand sensitivity to the preexisting relationship between the wrongdoer and her community, both from the perspective of the individual wrongdoer and from the perspective of the political community as a whole.

Likewise, it is true, as retributivists often emphasize, that social and political marginalization can render it more difficult (although not impossible) for wrongdoers to feel the force of moral reasons not to offend. For marginalized wrongdoers, it is often more difficult or burdensome to abide by the normative standards governing interpersonal and impersonal relationships, and the attitudes and actions appropriate to such contexts.\textsuperscript{182} However, despite the

\textsuperscript{180} It is important to note that, from the perspective of the community as a whole, the fact of marginalization is not only relevant to the appropriate expression of blame, but also calls for the community to attempt to repair the social circumstances that impair the preexisting relationship between the wrongdoer and her community. This point will be further discussed in what follows.

\textsuperscript{181} I am assuming, here, that the political community in question, while not perfectly just, meets some minimum threshold of legitimacy.

\textsuperscript{182} Consider, for example, the economically marginalized individual guilty of committing a property crime. It is plausible that it is more difficult or burdensome for a poor person (when compared with a wealthy person) to recognize the significance of violating certain normative standards, e.g. respecting the private property of other people. Conversely, it is plausible that it is less burdensome for a wealthy person (when compared with a poor person) to refrain from violating certain normative standards related to personal property. In either case, the relative burdensomeness of recognizing moral reasons does not render the
unfortunate increased burdensomeness of conforming to normative standards in
the case of marginalized agents, we should reject the argument that social and
political marginalization entirely vitiates moral responsibility. Nor, in my view,
does marginalization render blame unfair to the marginalized wrongdoer, given
that wrongdoing in such cases nevertheless impairs the relationships others can
have with the wrongdoer. The fact of marginalization, however, is relevant to
whether or not blame ought to be expressed, as well as the character and
severity of the expression of condemnation appropriate to any given context.

In my view, wrongdoers who are socially and politically marginalized
nevertheless ought to be considered morally responsible, i.e. marginalized
wrongdoers are appropriately subject to evaluation and criticism for their wrongful
attitudes and actions.\(^{183}\) I submit that it would be in an important sense
disrespectful toward marginalized wrongdoers to treat them as less than fully
rational and responsible moral agents. Marginalized wrongdoers are still rational
and responsible moral agents capable of responding to moral reasons, and guilty
of wrongdoing others. Furthermore, they still have obligations to others that stem
from their wrongdoing, i.e. the obligation to recognize that what they’ve done is
wrong, and the obligation to make amends to those whom they have wronged.\(^{184}\)

---

\(^{183}\) On one possible construal, such an evaluation might take the form of calling
upon marginalized wrongdoers to rationally justify their attitudes / actions to
others.

\(^{184}\) A restorative justice approach thus aims to reinforce the status of criminal
offenders, including marginalized offenders, as rational and responsible moral
agents capable of responding to moral reasons.
Nonetheless, from the perspective of the community as a whole, in my view, the particular mode of expressing the judgement that a marginalized wrongdoer is blameworthy ought to factor in some consideration of her marginalization. In the case of a relatively minor offence, but where the wrongdoer has been the victim of an extreme form of marginalization, we might judge it to be appropriate that our collective expression of the judgement that the wrongdoer is blameworthy remain merely formal, e.g. a criminal conviction. In cases involving more egregious forms of wrongdoing, but less significant marginalization, we might judge it to be appropriate that formal censure, e.g. a criminal conviction, should be combined with a deliberative process of shared moral inquiry, bringing together the wrongdoer, the victim(s), and directly affected members of the community. Of course, from a restorative justice perspective, the fact of marginalization is not only relevant to the appropriate collective expression of blame; a restorative justice approach demands that the community attempt to repair the social circumstances that impair the preexisting relationship between the wrongdoer and her community.\textsuperscript{185} Importantly, in any case, marginalized wrongdoers should not be “let off the hook.” Nor should the fact of marginalization eviscerate the legitimacy of blame expressed by the community. Our collective responsibility to blame wrongdoers in some appropriate way, while recognizing that blame can be expressed in a number of different ways, should

---

\textsuperscript{185} As argued above, criminal wrongdoing impairs the relationship the wrongdoer can have with others, especially the moral community. What I want to highlight here is that the social and political marginalization of the wrongdoer further impairs that relationship.
be exercised on the basis of judgements that are, at least in part, independent of the culpability of the wrongdoer.\textsuperscript{186}

Along similar lines, as Smith suggests, we should consider whether or not the wrongdoer in question is already remorseful, and thus already willing to make amends to those whom she has wronged. In such cases, again, while holding the remorseful wrongdoer accountable to the community is nevertheless appropriate, the particular mode of expressing the judgement that the wrongdoer is blameworthy and ought to be held accountable ought to factor in the judgement that she is already remorseful. We might in certain cases judge it to be inappropriately harsh to subject an already remorseful wrongdoer to hard treatment. Holding her accountable to the community and requiring that she live up to her obligations to others as a wrongdoer might be possible without the use of a process of state-imposed hard treatment. Again, merely formal public

\textsuperscript{186} One might object, here, by pointing out that factoring in considerations concerning the standing of the community in relation to marginalized wrongdoers would altogether undermine the authority of the state in the context of criminal justice to respond to criminal forms of wrongdoing. However, in my view, this problem also confronts idealist retributivist accounts that set the bar extremely high concerning the necessary preconditions for legitimate retributive punishment. See, for example, Murphy, “Marxism and Retribution.” In my view, in the context of criminal justice, we should be willing to invoke the familiar distinction between legitimate political authority and ideal circumstances of justice. In the context of a legitimate authority (but absent ideal circumstances of justice) we might think that the extent to which social culpability is apparent with respect to the standing of the community in relation to any given wrongdoer, such social culpability ought to be considered relevant to the appropriate response on the part of the community to criminal forms of wrongdoing perpetrated by marginalized offenders. It is thus the idealist version of retributivism that is more so vulnerable to the objection that such an account implies paralysis in the face of how we ought to respond to criminal forms of wrongdoing absent ideal circumstances of justice.
censure, or public censure plus a process of shared moral inquiry, might be deemed sufficient.

Of course, abandoning a retributivist approach to moral responsibility and blame, along with its implications for criminal punishment, involves abandoning a principle of strict proportionality in our responses to crime. This might strike some people, at least at first glance, as deeply counterintuitive. If the culpability of wrongdoers is not all that matters to holding wrongdoers accountable to the wider, normative community, then, at least at first glance, it will be more difficult to capture the firmly held intuition of many people that a principle of (strict) proportionality ought to rein in the permissible use of force in the context of criminal justice. However, if and how an approach to criminal justice ought to include a principle of (strict) proportionality will be determined by the specific theory of criminal justice in question. Thus far, all I have provided by way of argument are reasons to hold that a retributivist approach to moral responsibility and blame is not the only available approach to holding criminal wrongdoers accountable to the wider, normative community in the context of restorative justice. But, at the very least, we can be certain that abandoning retributivism requires that we relax the principle of proportionality that is at the heart of a retributivist approach to criminal justice.\textsuperscript{187}

\textsuperscript{187} This point will be further discussed in Chapter 6, where I address the objection that the account I defend (see Chapter 5) makes no room for a principle of proportionality. While I reject the notion (defended in certain restorative justice accounts) that proportional treatment plays no role in a restorative justice response to wrongdoing, I argue that the proportionality constraint at the heart of retributivist approaches must be relaxed, due to considerations independent of the culpability of offenders that are relevant to permissible punishment.
Finally, in my view, we should also include certain empirical, forward-looking considerations when determining the appropriate mode of expressing blame in the context of criminal justice. Among other considerations, we should consider the likely effects of our chosen mode of expression on wrongdoers, victims, and the wider normative community. For example, as Duff suggests, communicative punishment ought to aim at fostering moral recognition in wrongdoers. Therefore, the particular mode of expression ought to be sensitive to its likely effects on the wrongdoer.\textsuperscript{188} We should aim, as Duff suggests, at bringing the wrongdoer to recognize the moral gravity of her crimes.

Similarly, we should consider the likely effects of our chosen mode of expression on victims. In certain contexts, for example, it would be inappropriate to ask victims to sit down with their assailants to partake in a process of shared moral inquiry.\textsuperscript{189} In such cases, communicative punishment might be preferable to pursuing restorative justice via mere formal censure, or via a deliberative process of shared moral inquiry. However, in certain cases, the victim might view the opportunity to partake in a process of shared moral inquiry as an opportunity to further the valuable aim of moral recognition. As restorative justice theorists rightly point out, state-sponsored systems of criminal punishment are often experienced by the victims as deeply unsatisfying, insofar as the victims are excluded from the criminal justice process, in the sense of being denied the

\textsuperscript{188} Thus, in my view, empirical considerations concerning human psychology are relevant to holding wrongdoers accountable and pursuing the ends of restorative justice. This point will be taken up further in Chapters 5 and 6.

\textsuperscript{189} For example, Strang argues that power imbalances between offenders and victims points to a limitation of deliberative dialogue for the sake of restorative justice. See Strang, H., “Repair of Revenge,” pp.348-9.
opportunity to give a voice to their perspective on the crime and what ought to be done going forward to right any wrongs suffered.

Finally, we should consider the likely effects of our chosen mode of expression on the wider, normative community. In my view, as Duff suggests, the criminal law ought to aim at providing the public with morally relevant reasons not to offend. Nevertheless, crime prevention and societal protection are, in my view, legitimate aims to be pursued by a system of criminal law. Therefore, we should also consider the likely effects of our chosen mode of expression on the wider, normative community.¹⁹⁰

Before proceeding, it is worth responding to an objection concerning the relationship between criminal punishment and the ends of restorative justice specified by my account: an advocate of Duff’s view might object that criminal punishment is necessary to repairing the relationship between wrongdoers and their communities due to the *impersonal* nature of blame in such cases. Wrongdoers and their communities are related as *co-citizens*, rather than as friends, family members, acquaintances, etc. Therefore, one might conclude, punishment is necessary in order to repair the harm to such relationships, since, as a community we owe it to other members of the community (e.g. the victims, the family members of the victims, other law abiding citizens, etc.) to punish criminal wrongdoers.

In my view, however, in the spirit of Scanlon’s account of blame, whether we consider retributive punishment as necessary to repairing the impaired

¹⁹⁰ This argument will be further developed in Chapter 6.
relationships between co-citizens resulting from crime depends upon how we conceive of the relationship of “co-citizens.” While a retributivist construal of the appropriate response to damaged or impaired relationships between co-citizens is one available interpretation, it is not the only available interpretation. Another possible construal of the appropriate response to damaged or impaired relationships between co-citizens resulting from crime is that, as co-citizens, we ought to live up to what we owe to others after the fact of wrongdoing. Wrongdoers ought to recognize the wrongfulness of their conduct, and attempt to make appropriate amends. When relationships between co-citizens are impaired through wrongdoing, the appropriate response on the part of the community is an adjustment of the relationship between the community and the wrongdoer, which involves holding wrongdoers in some sense accountable to the community. As I argue in the next Chapter, one possible response on the part of the community is for the state to interfere in the lives of wrongdoers for the sake of persuading them to live up to what they owe to others as wrongdoers. An institution of criminal punishment that metes out forceful public censure is but one available institutional response among others. As a community, how we decide to respond to criminal wrongdoing will depend upon how we conceive of our relationships with each other as co-citizens, and how we conceive of the role for blame, apology, and the restoration of relationships appropriate to that context.

4.6 – Conclusion
In this Chapter, I argued against the two dominant views of restorative justice. On the one hand, the informal model of restorative justice falls short in at least one important respect: the prescribed process of deliberative dialogue, understood as a private process of shared moral inquiry, sets to one side the important consideration that the wider, normative community has an obligation to hold wrongdoers accountable for the commission of public wrongs. On the other hand, I argued that restorative retributivism presupposes a contestable retributive theory of moral responsibility and blame. Once we distinguish rigorously between judgements of blameworthiness, on the one hand, and the appropriate expression of such judgements, on the other, we have reasons to pursue a more nuanced approach to holding wrongdoers accountable to the wider, normative community. The appropriate mode of expressing the judgement that the criminal wrongdoer is blameworthy and hence ought to be held accountable to the community ought to factor in considerations that are independent of the culpability of the wrongdoer. Thus, on this view, the retributivist approach to blaming offenders, along with its commitment to strict proportionality, are rejected.

Thus far, however, I have yet to put forward an alternative theory of criminal punishment and restorative justice. Retributivism, despite its weaknesses, is clear about the terms of the wrongdoer’s liability to suffer punishment, i.e. punishment is something that wrongdoers (in some sense) deserve to suffer. In rejecting the restorative retribution view, therefore, it is necessary to provide a different account of why it ought to be considered
permissible to subject wrongdoers, at least in certain cases, to hard treatment for
the sake of pursuing the ends of restorative justice. Providing such an argument
is the purpose of the next Chapter. In what follows, I argue that the duties of
wrongdoers as wrongdoers provide the normative framework for permissible
criminal punishment, as one institutional possibility among others for holding
wrongdoers accountable to the community.
Chapter 5 – A restorative justice theory of punishment

5.1 – Introduction

In the previous Chapter, I explored the merits of the restorative justice alternative to the problem of criminal wrongdoing. I argued that we should reject the two dominant approaches to restorative justice. On the one hand, we should reject the view that criminal punishment is necessarily at odds with pursuing the ends of restorative justice. On the other hand, we should also reject the view that restorative justice requires retributive punishment. I argued that the former view, the informal model of restorative justice, emphasizes plausible forward-looking aims, including fostering moral recognition in wrongdoers, and reparations between wrongdoers, their victim(s), and affected members of the normative community, while at the same treating wrongdoers with the respect they are owed as rational and responsible moral agents. I argued that the latter view, restorative retributivism, draws our attention to a significant weakness associated with the informal model: the wider, normative community has an obligation to hold wrongdoers accountable for the commission of public wrongs. A state-sponsored response to criminal wrongdoing is therefore an important aspect of pursuing the ends of restorative justice. Subsequently, by appealing to a non-retributivist theory of moral responsibility and blame, I developed an alternative, non-retributivist approach to criminal responsibility and restorative justice.
In this Chapter, I defend a non-retributive restorative justice theory of punishment grounded in the duties of offenders. I argue that criminal punishment ought to be considered permissible, albeit in a way that is highly restricted or constrained, for the sake of achieving certain valuable ends. More specifically, I argue that an institution of criminal punishment, if and when such an institution ought to be considered permissible, ought to be considered permissible for the sake of pursuing what I call the ends of restorative justice: criminal punishment must aim at fostering moral recognition in wrongdoers, toward the further end of reconciling wrongdoers with their communities. I argue that a famous, traditionally retributivist objection to this sort of view, i.e. that any fundamentally forward-looking justification for punishing wrongdoers will ultimately fail to respect the status of wrongdoers as rational and responsible moral agents, can be overcome by appealing to a plausible substantive account of what wrongdoers owe to their victims and others in the community as rights-bearers.\footnote{For a version of this view, see Duff, \textit{Punishment}. As discussed in Chapter 4, while I reject the retributivist dimension of Duff's account – i.e. the retributive punishment is a necessary condition for the pursuit of restorative justice – in what follows I rely heavily on Duff’s view surrounding the relevant obligations of wrongdoers as wrongdoers, i.e. the duties of moral recognition specified by his account. See, for example, Duff, “Restorative Punishment,” pp. 372-3.} Wrongdoers, I argue, incur second-order enforceable obligations of moral recognition\footnote{\textit{Ibid.}} (section 3). Such duties generate limited permissions to punish wrongdoers for the sake of pursuing the ends of restorative justice (section 4). In this sense, a restorative justice theory of punishment is able to account for both the moral agency and the basic rights of wrongdoers: focusing on the secondary...
obligations of offenders captures an attractive ingredient typically associated with retributivism, namely, that criminal punishment ought to address wrongdoers as wrongdoers. A restorative justice approach grounded in the duties of offenders thus aims to treat wrongdoers as rational and responsible moral agents who possess basic rights, while also orienting the justification for criminal punishment around valuable forward-looking aims, i.e., the ends of restorative justice. Subsequently, I highlight some limitations associated with the restorative justice account (section 5), and conclude by situating my view in the context of some familiar retributivist themes (section 6).

5.2 – Restorative justice and the duties of offenders

As discussed in the previous Chapter, restorative justice is about healing or repairing the damage done to the community resulting from criminal forms of wrongdoing. One plausible way of construing the damage done to the community resulting from crime is in terms of the damage done to the normative relationships between wrongdoers and their communities (and between wrongdoers and their victims as co-members of a shared community). A reparative or restorative approach to criminal justice, therefore, aims to heal or restore the normative relationships between wrongdoers and victims as co-members of a shared, normative community.

Along these lines, Duff argues that communicative punishment, in order to be permissible, must be conducive to fostering moral recognition in wrongdoers,

\[\text{\cite{193} See Chapter 4, section 2.}\]
toward the further end of reparations between wrongdoers and their communities. As previously discussed, for my purposes I refer to these ends collectively as *the ends of restorative justice*. Recall that, for Duff,

> Communicative punishment … should be understood and justified as a communicative, penitential process that aims to persuade offenders to recognize and repent the wrongs they have done, to reform themselves, and so to reconcile themselves with those they have wronged.  

As I argued in Chapter 4, however, Duff’s approach constitutes a form of restorative retributivism, in the sense that he views retributive punishment as necessary to pursuing the ends of restorative justice. What renders wrongdoers liable to hard treatment for the sake of pursuing such ends, on Duff’s account, is that hard treatment is what the wrongdoer “deserves for her crime.” For Duff, censuring wrongdoers via hard treatment is the intrinsically appropriate response to crime on the part of the normative community. As I argued in Chapter 4, however, Duff’s account rests on a strongly retributivist account of moral responsibility and blame. In my view, by contrast, if we distinguish carefully between the judgement of blameworthiness on the one hand, and the appropriate expression of blame on the part of the normative community on the other, then it is possible to view the relationship between criminal punishment and restorative justice in non-retributivist terms; the appropriate mode of expressing the judgement that the criminal wrongdoer is blameworthy and hence ought to be held accountable to the community ought to include considerations that are independent of the culpability of the wrongdoer.

---

195 See Duff, “Restorative Punishment,” p.380. Also see Chapter 4, section 4.
However, this strategy engenders a new problem: the retributivist strategy is clear about the terms of the wrongdoer’s liability to suffer punishment for the sake of restorative justice; for the restorative retributivist, wrongdoers deserve to suffer punishment as the intrinsically appropriate response to their wrongdoing. According to the non-retributivist strategy I developed in the previous Chapter, however, a different account is needed concerning what grounds the liability of wrongdoers to suffer hard treatment for the sake of pursuing the ends of restorative justice. I described this problem in Chapter 2 as the problem of the basic rights of wrongdoers: respecting the basic rights of wrongdoers requires some form of reconciliation between the forward-looking aims of an institution of criminal justice and the basic rights of wrongdoers against being used “as a means” in certain ways. In my view, as I argue in what follows, the defender of a restorative justice approach can and ought to appeal to the secondary duties of wrongdoers. My aim is to articulate and defend a non-retributivist restorative justice theory of punishment, grounded in the secondary duties of wrongdoers to recognize and express recognition of the wrongness of their criminal acts.

5.3 – Moral recognition and the duties of offenders

Recently, philosophers have turned to a relatively new set of normative considerations when thinking about an institution of criminal punishment and its appropriate justification. Rather than focusing primarily on what we ought to do to wrongdoers, the focus has shifted to what wrongdoers themselves ought to do, given that they are guilty of wronging others. In searching for a justification for
criminal punishment, philosophers are now turning their attention to the obligations of wrongdoers as wrongdoers.\textsuperscript{196}

The view I aim to defend falls outside of the retributivist camp, broadly understood, since it justifies punishment not with reference to its intrinsic value (or, in other words, not with reference to the desert of offenders). On the contrary, in my view, the value of punishment derives from the valuable ends that might be achieved via an institution of criminal punishment, namely, the ends of restorative justice. Importantly, due to the nature of such ends, I will argue, the scope of permissible punishment ought to be considered extremely narrow; so narrow, I will argue, that such an account calls into question the permissibility of most, if not all, actually existing institutions of criminal punishment.\textsuperscript{197}

A famous traditionally retributive objection to forward-looking theories of punishment is that they tend to fail to respect the moral status of wrongdoers as rational and responsible agents, and as individuals who possess basic rights against being used “as a means” in certain ways. Therefore, if an institution of criminal punishment is to be justified with reference to its forward-looking value in bringing about the ends of restorative justice, i.e. restoring offenders, and repairing the normative relationships impaired by criminal forms of wrongdoing,

\textsuperscript{196} See, for example, Tadros, V., \textit{The Ends of Harm: The Moral Foundations of Criminal Law} (Oxford: Oxford University Press, 2011). To my knowledge, Tadros’s view represents the first attempt to ground a comprehensive theory of punishment in the duties specific to wrongdoers as wrongdoers. I disagree with Tadros’s view surrounding what he considers to be the relevant duties of offenders, and surrounding his contention that the relevant duties of offenders justify a deterrence approach to criminal punishment, but this is beyond the scope of my discussion here.

\textsuperscript{197} I return to this point in section 5.
so too must such an account face up to the objection that criminal punishment, in
order to be morally legitimate, must address wrongdoers as rational and
responsible moral agents who possess basic rights. This is where I want to
suggest that a restorative justice theory of punishment can and ought to appeal
to a substantive account of what wrongdoers owe to their victims and others in
the community as rights-bearers.

On one plausible version of this view, wrongdoers incur second-order
duties of moral recognition.198 I consider a duty to be “second-order” when it
follows from, or is a normative implication of, a failure to live up to a first-order
obligation. So, for example, if I culpably fail in my duty to fulfill my promise to you,
say, to pick you up at the airport, it is plausible that I, in virtue of wronging you in
this way, incur further obligations. We might think that I owe you an apology, and
that I ought to compensate you for your loss in whatever way I can, perhaps by
assuming the cost of your subsequent cab ride. In the case of criminal forms of
wrongdoing, (or, put in strictly moral terms, in the case of public instances of
moral wrongdoing), by contrast, we might think that the relevant second-order
obligations of wrongdoers are to a greater degree stringent and onerous. This is
because, in my view, criminal forms of wrongdoing arguably involve a particularly
pernicious attitudinal dimension, i.e. disrespect for certain specified normative

---

198 See, for example, Duff, *Punishment*, p.95. Also see Duff, “Restorative
Punishment,” pp.372-3. For a discussion of the differences between my view and
Duff’s account, see Chapter 4, especially sections 4-5.
standards, ultimately resulting in serious injury to the relationships the wrongdoer can have with others.199

My view involves a commitment to distinguishing, in some concrete sense, between moral wrongs in general, and forms of moral wrongdoing that are appropriately subject to criminalization. Criminal wrongdoing, in my view, must involve a pernicious attitudinal dimension that impairs the relationships that the wrongdoer can have with others. However, importantly, attitudinal impairment though necessary is not a sufficient condition for a moral wrong to be appropriately subject to criminalization. There are pernicious attitudes that impair the relationships that wrongdoers can have with others that would not be appropriate to criminalize. For example, a form of wrongdoing might not be appropriately the concern of the community as a whole. For example, it might be morally wrong for me to taunt my friend in public about his failed personal projects. However, most would agree that such a form of wrongdoing ought to be considered between my friend and me, so to speak (rather than between my friend, me, and the moral community as a whole). In such a case, most would agree, it is not the appropriate role for the community as a whole to get involved. Other forms of wrongdoing, by contrast, typically are considered public wrongs, in the sense that they are the appropriate concern of the community as a whole.

199 In my view, for a wrongdoer to be liable to coercive forms of treatment on the basis of her criminal misconduct, the form of wrongdoing in question must be a “public wrong,” in the sense of being the appropriate concern of the moral community as a whole. See Chapter 4, sections 4-5. This point will also be further discussed in what follows. For a discussion of criminal wrongs as public wrongs, see Duff, Answering for Crime, p.140-146.
As discussed in the previous Chapter, domestic violence, for example, appropriately concerns the community as a whole.\textsuperscript{200}

Return now to the topic of the secondary duties of offenders. As I suggested above, we might think that the secondary duties of offenders render them liable to coercive forms of treatment, in order to pursue through institutionalized coercion (within certain limits) the ends of restorative justice. The argument is as follows: under normal conditions, everyone has a right to be treated as a human being, and thus as a being worthy of moral consideration. Correspondingly, people have duties under normal conditions to refrain from acting in ways that undermine or ignore the moral status of others.\textsuperscript{201} For my purposes, I refer to such duties as primary duties of moral recognition. However, wrongdoing is a special case, in that wrongdoers have demonstrated via the commission of wrong acts a lack of concern for the moral status of others.

\textsuperscript{200} See Chapter 4 sections 4-5. However, to complicate matters further, as Duff argues, the public dimension of the criminal wrong does not ground, but rather is a condition for, the appropriateness of its criminalization. Therefore, on this view, while all appropriately criminalized forms of wrongdoing are public in the requisite sense, not all public wrongs necessarily ought to be criminalized. What grounds the scope of the criminal law in any given context ultimately flows from a deeper set of normative considerations concerning the just political society, and the core values deemed to be appropriately the business of the just society as a whole so specified. Such questions are, however, subject to reasonable disagreement. Duff writes, “A justification of criminalisation will need to begin by specifying some value(s) that can be claimed to be public, as part of the polity’s self-definition; show how the conduct in question violates that value or threatens the goods that it protects; and argue that the violation or threat is such as to require or demand public condemnation.” Duff, \textit{Answering for Crime}, p.143.

\textsuperscript{201} Criminal forms of wrongdoing do not necessarily involve actually setting back the interests of other people. Consider the case of drunk driving: the act of driving while under the influence of alcohol wrongs other people, not only in the case where it actually results in injury or harm to others; it is wrong purely in the sense that it involves a lack of regard for the moral status of others. For this point see Kumar, “Who Can Be Wronged?” p.103.
Therefore, the argument is that secondary to a greater degree onerous duties kick in after the fact of wrongdoing, to recognize and express recognition of the moral status of others, including the victim(s).

Along these lines, it is necessary to distinguish between two senses in which wrongdoers have an obligation (at the secondary level) to recognize the moral status of their victims and others: (1) the attitudinal sense and (2) the expressive or communicative sense. First consider (1). The attitudinal dimension of the secondary duty to recognize the moral status of others involves a process of moral reform on the part of the offender.\textsuperscript{202} To the extent that an instance of criminal wrongdoing involves blatant or passive disregard for the moral status of others, wrongdoers have a duty, especially to those whom they have wronged, but also to the wider, normative community, to adopt attitudes (understood as intentions, beliefs, desires, etc.) that do not involve blatant or passive moral disregard for others.\textsuperscript{203} On this point, in the context of restoring the normative relationship between offenders and their communities, Duff argues, “Recognition of the wrong as a wrong is clearly an indispensable first step. It is something that is owed to the victim by others in general … but especially by the wrongdoer.”\textsuperscript{204}


\textsuperscript{203} In the interpersonal context, it is a famous point in the philosophy of moral responsibility that the attribution of blameworthiness to some agent involves attributing to her certain attitudes, in particular attitudes that impair her relationships with others. See, for example, Scanlon, \textit{Moral Dimensions}.

\textsuperscript{204} See Duff, “Restorative Punishment,” p.372. As stated above, I owe my account of the duties of offenders to Duff’s communicative theory of punishment. As will be discussed in what follows, however, making the duties of offenders rather than the desert of offenders central to my account (in the sense of
Now consider (2). Moral reform on the part of the offender, in the sense of fulfilling her secondary obligation of moral recognition in the attitudinal sense, is arguably not enough for her to fully live up to what she owes to others.

Wrongdoers arguably also owe it to others to communicate genuine apology, understood as an expression of genuine recognition of the moral status of others, especially the victims, and remorse for the instance(s) of wrongdoing involved.  

Making a similar point, again in the context of restoring the normative relationship between offenders and their communities, Duff argues, “What I owe to those I have wronged, and what others may try to persuade me to if I need persuasion, is … a repentant recognition of the wrong I have done – a repentance that apology can … express.”

Under normal conditions, therefore, recognizing the moral status of others as human beings requires that we refrain from violating or ignoring the basic moral status of others in our actions and inactions. In the special case of wrongdoing, however, since the wrongdoer has already demonstrated, via the commission of wrong acts, blatant or passive disregard for the basic moral status of others, i.e., failed in her primary duties, the duty to recognize the moral status providing the ground for the liability of wrongdoers to suffer punishment) leads us in the direction of a substantively revised version of the communicative theory of punishment; communicative punishment, in my view, is not something that wrongdoers deserve to suffer per se, but rather something that is, in a highly restricted sense, permitted for the sake of pursuing the ends of restorative justice. This argument will be further elaborated in what follows. Therefore, in my view, making the duties of offenders central to the account leads us in the direction of a non-retributivist construal of the communicative account. Cf. Chapter 4, section 5.

For a similar claim, see Duff, *Punishment*, p.95.

of others requires that the wrongdoer do more than merely refrain from acting in ways that undermine or ignore the basic moral status of others. For the wrongdoer, the more onerous secondary obligation of moral recognition in the expressive sense involves the burden of communicating recognition, in the form of a meaningful apology, to her victims and others. In this way, the positive duty of communicating recognition in the special case of wrongdoing takes the place of the merely negative duty associated with refraining from acting in ways that demonstrate indifference or disregard for the basic moral status of others under normal conditions.

Importantly, however, given egregious instances of wrongdoing, mere verbal apology is arguably not in itself enough to discharge the secondary duty of recognition in the expressive sense. As Duff argues, “Sometimes … a (mere) apology is not enough. If I have done a serious wrong to another person, I cannot expect to settle or resolve the matter merely by [verbally] apologizing to him: something more than that is due to him and from me.”207 As Duff suggests, it is plausible that wrongdoers owe something to their victims and others (as co-members of the normative community) that goes beyond mere verbal apology, i.e. wrongdoers owe it to others to make a meaningful attempt to demonstrate their understanding, recognition, and ultimately remorse for what they have done. It is plausible, as Duff argues, that one of the central purposes of communicative punishment is to add moral weight to the apology of the offender, allowing her

207 Ibid.
the opportunity to fulfill her obligations of recognition to those whom she has wronged, and the moral community more generally.\textsuperscript{208}

Understood in this way, the secondary obligations of offenders explain why failing in one’s secondary obligation of moral recognition constitutes a further, distinct form of moral wrongdoing directed at victims and others in the community. When wrongdoers fail to reform themselves, or worse, express indifference or other disrespectful attitudes in relation to their wrongdoing and those whom they have wronged, they ought to be considered guilty of a further, distinct form of moral wrongdoing. For example, were I to violently assault you and subsequently fail to reform myself, or worse, express an attitude of indifference regarding my assault upon you (say, by smirking in your direction, or perhaps by stating explicitly, “I don’t care that I’m guilty of assaulting you, in fact I...

\textsuperscript{208} See, for example, Duff, \textit{Punishment}, p.109. This point will be further discussed in what follows. I should note that the two senses of the secondary obligation of moral recognition described above are closely related. For this reason, I describe them as “senses” of having an obligation, rather than as distinct obligations. For example, discharging (1) is a necessary condition for discharging (2). If I fail to reform myself, i.e., to recognize the normative standards my attitudes / actions have violated, and hence the wrongness of my conduct and the importance of the moral status of my victims and others, then I am not in a position, morally speaking, to meaningfully demonstrate my remorse and communicate genuine apology. This is an important sense in which my view represents a departure from Duff’s retributivist account that finds value in merely “formal” or “ritualistic” forms of apology. See Duff, “Restorative Punishment,” p.378. For Duff, “…there can be value … even in ritual or formal apology; and we can hope that the experience of making even a demanded, non-voluntary apology might help bring the offender to a clearer grasp of the character and implications of what he has done.” In my view, in the case wherein apologetic action is forced upon an unrepentant offender, we as a community fail to promote the ends of restorative justice, although (at least in many cases) such a failed attempt may nevertheless be considered permissible. I am skeptical of the more specific (retributivist) claim that such a failed attempt is in itself (symbolically or otherwise) valuable, or succeeds in repairing the normative relationship between the wrongdoer, her victims, and her community. Cf. Bennett, \textit{The Apology Ritual}. 123
enjoyed it,” etc.), I thereby continue to wrong you in an importantly distinct, additional way besides my original assault upon you, because I am also failing in my secondary obligation of moral recognition.

In sum, I contend that an institution of criminal punishment ought to be considered valuable purely for the sake of pursuing the ends of restorative justice. Therefore, if an institution of criminal punishment is to be justified, in my view, it must be justified with reference to what wrongdoers owe to others, and the value of an institution of criminal punishment in bringing about (or, perhaps more accurately, approximating in a second best way) the relevant ends related to what wrongdoers owe to others, i.e. the ends of restorative justice. On this view, therefore, there is no essential relationship between an institution of criminal punishment and the ends of restorative justice.\textsuperscript{209}

\textbf{5.4 – Duties of moral recognition and the limits of enforcement}

It is, at least in principle, possible – although it is perhaps in many cases very unlikely – that wrongdoers might live up to what they owe to others as co-members of a shared, normative community, without suffering the hard treatment imposed by a system of criminal punishment. However, a plausible theory of criminal punishment must address, as a matter of justification, the coercive aspect of a state-sponsored system of criminal punishment. We might frame the matter in terms of an objection: an opponent might object to my argument thus far by pointing out that a system of criminal punishment is a coercive institution,

\textsuperscript{209} Cf. Chapter 4, sections 4-5.
i.e. it is something that is almost always forced upon wrongdoers. Along these lines, one might object that even if it is true that wrongdoers incur certain secondary obligations, it remains to be established that punishment for the sake of restorative justice is something that we, as a community, can legitimately force upon wrongdoers. Perhaps the secondary obligations of wrongdoers are not the sort of obligations, it might be argued, that we can – or that we would want to, even if we could – compel wrongdoers into fulfilling. In my view, however, secondary obligations of recognition constitute examples of enforceable obligations, rather than merely something wrongdoers ought to do (although it will be critical to my view in what follows to address how and to what extent such obligations ought to be considered enforceable).

There are good reasons to think that enforcing wrongdoers’ secondary duties of moral recognition ought to be considered as among the legitimate aims of a coercive system of criminal law.\textsuperscript{210} Coercion is justified where fulfilling the duty in question is critical to the relevant basic rights or vital interests of others in the community, and where the duty-bearer has demonstrated an unwillingness to fulfill her obligations without interference.\textsuperscript{211} It is an extremely weighty concern, in my view, that members of the community respect the status of other members as morally considerable individuals. Fulfilling our obligations toward others to refrain from acting in ways that demonstrate utter disregard for their moral status is an

\textsuperscript{210} Here I agree with Tadros. See Tadros, \textit{The Ends of Harm}, pp.3-4.
\textsuperscript{211} For a slightly different argument concerning what justifies considering a duty to be an \textit{enforceable} duty, see Tadros, \textit{The Ends of Harm}, p.189.
extremely important moral responsibility, i.e. it is something that we owe to each other.

However, we would not be able to (nor would we want to, even if we were able to) compel everyone into fulfilling primary obligations of recognition at all times. To even attempt to do so would be far too intrusive, and involve far too much coercion, and would, for these reasons, be too deeply in tension with familiar liberal principles associated with freedom from state interference in our personal lives. In the special case of wrongdoing, however, when the wrongdoer has demonstrated, through her actions or inactions, disregard for the moral status of others, then compelling her to fulfill her secondary (to a greater degree onerous) duties of moral recognition becomes all the more important, and would involve, overall, much less intrusiveness. If people are unwilling to fulfill a certain set of their basic obligations toward others at the primary level, then the argument is that the state ought to intervene in order to attempt to convince wrongdoers, through the use of force, to fulfill a certain set of their basic obligations toward others at the secondary level. On these grounds, we have good reasons to set up a coercive set of institutions to ensure that we don’t fail each other in this extremely important matter.

Before moving on, it is necessary to make clear the extent to which secondary obligations of moral recognition ought to be considered enforceable obligations; there is arguably an important tension between having a certain obligation (especially with respect to certain types obligations), on the one hand, and being compelled into fulfilling such an obligation, on the other. Again, it is
possible to frame the point as an objection: one might object to my argument thus far by pointing out that moral recognition, unlike, for example, protection from future harm, is something that cannot be coerced. It might rightly be added to this complaint that even if it were possible to coercively extract moral recognition from wrongdoers by force, it would be undesirable to do so for moral reasons. Therefore, one might conclude, it would be misguided to use force to attempt to compel wrongdoers to recognize and express recognition of the moral status of others.

This objection, in my view, raises an extremely important point: even if it were possible to extract moral recognition from wrongdoers by sheer force, doing so (at the secondary level) would be undesirable, since the point of (secondary) moral recognition is that it is an expression of the agency of the wrongdoer, as the particular agent responsible for wronging others. Consider the following case:

*Manipulation* – Norman is guilty of violently assaulting Martha. In response, as a form of punishment, the state injects Norman with a psychoactive drug that forces him henceforth to think and behave in ways consistent with recognition and respect for the moral status of Martha and others.

The problem with the treatment of Norman in *Manipulation*, it seems to me, is that it allows no room for him as a moral agent to fulfill his secondary obligation of recognition toward Martha and others. Injecting him with the drug merely manipulates his body in a way that resembles moral recognition toward others. However, when we demand as a community that wrongdoers take responsibility for their wrongful actions, it is something that we demand of wrongdoers as rational and responsible moral agents. The punishment of Norman in
**Manipulation** therefore misses the point, i.e. that Norman, as a moral agent, ought to live up to what he owes to others in the community.

Thus moral recognition is not something that we would want to (even if we could) extract from wrongdoers by sheer force. However, it is nevertheless possible that a coercive set of institutions might be designed in ways that aim, through the use of force, to foster recognition and elicit remorse from wrongdoers, in ways that also respect the status of wrongdoers as rational and responsible moral agents, i.e. without attempting to extract recognition and remorse from them by sheer force, as in **Manipulation**. Consider the following case:

*Incarceration*: Dorothy is guilty of violently assaulting Ethan. In response, the state sentences Dorothy to two years in prison.

Compare, for a moment, the treatment of Dorothy in *Incarceration* to the treatment of Norman in *Manipulation*. Norman is psychologically and behaviorally coerced, and in that sense he is manipulated; the treatment of Norman in *Manipulation* highjacks his moral agency. The treatment of Dorothy, by contrast, coerces her only to the extent that she is incarcerated. During her incarceration (and after) Dorothy is left to think and act freely as a moral agent.²

In this way, an institution of criminal punishment may be carefully designed in order to facilitate and enable the discharge of wrongdoers’

---

²Dorothy is left free to act, of course, only within the circumscribed context of confinement. It remains an open question whether the treatment of Dorothy in this case is permissible. Instead, what I want to highlight is that the treatment of Dorothy in *Incarceration* radically curtails her freedom, although *in a different way* than the treatment of Charlie in *Manipulation*, i.e. it leaves her status as a moral agent intact.
secondary duties of recognition, especially in situations where wrongdoers display an initial reluctance to fulfill their secondary duties on their own. On this view, the hard treatment aspect of criminal punishment ought to be designed so as to bring about moral recognition in wrongdoers, but never in ways that wholly bypass or undermine the moral agency of wrongdoers, as in *Manipulation*. Criminal punishment that is justified on the grounds of the duties of offenders, therefore, ought to be designed not to force wrongdoers into fulfilling their obligations of recognition per se, but rather to attempt, through the use of force, to convince or persuade wrongdoers, as moral agents, to themselves fulfill their obligations to others (in both the attitudinal and expressive senses). Justifiable criminal punishment, therefore, ought to aim at both compelling and enabling wrongdoers to fulfill their secondary obligations, toward the further end of reconciling themselves with their victims and others as co-members of a shared, normative community.

In sum: the duties of offenders create limited permissions for a third party, e.g. the state, on behalf of the normative community, to interfere in the lives of wrongdoers for the sake of convincing them to live up to what they owe to others, i.e. in order to convince them to fulfill their secondary obligations of moral recognition. On this view, the hard treatment aspect of criminal punishment ought to be designed so as to foster moral recognition in wrongdoers, toward the further ends of restorative justice. The value of an institution of criminal punishment, therefore, derives from the more basic values associated with the ends of restorative justice, i.e. the ends of fostering moral recognition in
wrongdoers, and reparations between wrongdoers and their communities. Critically, the achievement of recognition on the part of wrongdoers, and the discharge of their secondary duties (in both relevant senses) must ultimately flow from their own moral agency. The apparent tension between a coercive institution aimed at fostering moral recognition in wrongdoers, and the agency of wrongdoers as a necessary condition for the achievement of genuine moral recognition, is resolved by attending to the limited character of the coercive treatment justified by the secondary obligations of offenders. In order to be permissible, criminal punishment must be constrained in ways that leave intact the moral agency of wrongdoers, allowing them substantive opportunities to reform themselves, and to make amends with their victims and others as co-members of a shared normative community.\textsuperscript{213}

5.5 – Limits of the restorative justice account

The justification for criminal punishment supplied by my argument above is limited in several important respects. For my immediate purposes, three such limitations are especially noteworthy:

\textsuperscript{213} I should add that, in my view, when a form of interference fails to achieve its desired aim, i.e. when coercive interference fails to convince the wrongdoer to live up to what she owes to others in the requisite sense, the interference remains permissible, although it fails to achieve the valuable aim that justifies its imposition. Cf. note 208 above. Furthermore, the use of coercion for the sake of persuading wrongdoers to live up to what they owe to others ought to be considered constrained by a principle of proportionality, although, pursuant to the discussion in Chapter 4 regarding retributivism and proportionality, in my view, the use of coercion for the sake of pursuing the ends of restorative justice should not be constrained by a principle of strict proportionality. This argument will be taken up further in the next Chapter (Chapter 6).
(1) the first important sense in which the argument above is limited or restricted relates to a theoretical distinction between punishment for the sake of restorative justice, on the one hand, and other modes of non-punitive treatment aimed at the same ends, i.e. the ends of restorative justice, on the other. Even if wrongdoers incur certain duties, we should resist jumping to the conclusion that a system of coercive punishment is, in principle, the best or only way to interfere in the lives of wrongdoers for the sake of persuading them, through the use of force, to discharge their secondary obligations. It might turn out that, in some given context, a system of criminal punishment is indeed an efficient (or the most efficient) means for the sake of pursuing such ends. However, it also might turn out otherwise, i.e., that other forms of coercive treatment are more conducive to fostering recognition in wrongdoers, and so more conducive to allowing for wrongdoers to discharge their relevant secondary obligations. Even if we accept that the duties of offenders supply a plausible, substantive argument establishing the permissibility of punishment for the sake of pursuing the ends of restorative justice, a further argument is necessary in order to show that a certain policy ought to be adopted by the state for the sake of pursuing the ends of restorative justice (via a system of criminal punishment, or via some other coercive system), all things considered. It is conceptually possible, in my view, that some other institution involving non-punitive coercive measures, e.g. merely formal, public censure, or formal public censure plus a mandatory process of shared moral inquiry,²¹⁴ might be more conducive to achieving the ends of restorative justice,

²¹⁴ Cf. Chapter 4, section 5.
given all relevant considerations. We should not assume, \textit{a priori}, that a system of criminal punishment is always going to be the most efficacious in bringing about the ends of restorative justice (even though, \textit{a posteriori}, in certain cases, it might turn out to be true).

Recall that commentators have criticized Duff on the basis that there is arguably no necessary relationship between the aim of censuring wrongdoers and a system of criminal punishment. As Feinberg argues, public censure might take a merely formal or symbolic form, such as a public admonishment of some sort.\textsuperscript{215} In my view, according to a restorative justice approach, we should accept that there is no necessary connection between holding wrongdoers accountable, in the sense of blaming them for their wrongful conduct, and an institution of criminal punishment.\textsuperscript{216} However, we should consider an institution of criminal punishment as one institutional possibility among others for holding wrongdoers accountable to the community, and persuading them, through the use of force, to live up to what they owe to others.

(2) A second important sense in which the argument above is limited or restricted is worth outlining: if a system of criminal punishment is to be justified, then a considerable degree of empirical investigation into the appropriate \textit{forms} of punishment (as well as alternative modes of coercive criminal treatment)

\textsuperscript{215} See Feinberg, “The Expressive Function of Punishment” p.423. Also see Honderich, \textit{Punishment: The Supposed Justifications}, p.182. Recall that Honderich argues, “If your aim is to \textit{tell somebody something}, is it necessary to do so by [for example] putting him jail for 20 years?”

\textsuperscript{216} This gives rise to an important objection to my account on the part of advocates of Duff’s view, namely, that retributive punishment ought to be considered essential to pursuing the ends of restorative justice. For my response to this objection, see Chapter 4 section 5.
would be necessary, in order to supply good reasons to think that this or that form of punishment (or this or that form of coercive, criminal treatment) is more or less conducive to fostering moral recognition in wrongdoers. It is likely that the forms of punishment rendered permissible by such an account, given the restrictions set out by the secondary duties of offenders, will be extremely limited in scope; forms of punishment that fail to include genuine, substantive opportunities for wrongdoers to discharge their obligations specific to them as wrongdoers cannot be justified.

(3) A third important sense in which my argument is limited or restricted relates to the distinction between theory and practice, and is closely related to the previous limitation (2) outlined above: my argument puts forward an ambitious normative ideal of criminal punishment and its appropriate justification, rather than a defense of the real-world status quo, or, for that matter, any actually existing system of criminal punishment. My argument should not be misconstrued, therefore, as providing a set of reasons in support of the claim that, for example, the institution of criminal punishment in the United States, Canada, or anywhere else, ought to be considered morally permissible. On the contrary, due to the normatively ambitious nature of the account I defend – i.e. that morally legitimate criminal punishment must involve genuine, substantive opportunities for wrongdoers to discharge the obligations that are specific to them as wrongdoers – it is unlikely that any presently existing system of criminal punishment can be said to live up to such an ideal. The real world limitations of criminal punishment, however, should not be taken as reasons to reject the ideal.
Rather, the ideal set out by my account should serve as a catalyst for considering the matter of reforming criminal justice institutions, prompting a cluster of important questions that would emerge from a central question, “How is it possible to get from here to there?” Such questions, it is important to note, will undoubtedly prove to be difficult, complicated, and controversial, but also highly important. 

5.6 – Restorative justice and retributivist themes

The argument above supplies a normative framework, grounded in the duties of offenders, for the permissibility of criminal punishment that aims at the ends of restorative justice. Before moving on, it is worth situating the restorative justice account in context, in particular, in relation to the retributivist themes discussed at length in Chapter 3. While my account is non-retributivist in character, in the sense that it finds no intrinsic value in an institution of criminal punishment, it is nevertheless able to capture certain attractive elements traditionally associated with retributivism. Most straightforwardly, the restorative justice account is able to capture the idea that criminal justice institutions ought to treat wrongdoers as wrongdoers. In my view, our responses as a community to criminal forms of wrongdoing should recognize the wrongdoer as a wrongdoer, i.e. as an agent who is culpably responsible for wronging others, and hence is appropriately the object of moral criticism and condemnation. However, as I

Footnote: For a similar yet much more detailed discussion of ideal theory versus the moral demands of real-world criminal punishment, see Duff, *Punishment*, Ch. 5.
argued in Chapter 4, how we choose to express the judgement that wrongdoers are culpably responsible, and hence ought to be held accountable to the community, ought to factor in various considerations, including considerations independent of the culpability of the wrongdoer.

We are now in a position to factor in the consideration of to what extent hard treatment should play a role in the mode of expressing the judgement of blame: the duties of offenders render permissible but also ought to constrain our use of hard treatment in the context of criminal justice. Criminal punishment, as a mode of forceful, public censure, should aim to persuade wrongdoers to live up to what they owe to others in the attitudinal and expressive senses. Therefore, the forms of punishment we impose upon wrongdoers should address wrongdoers as rational and responsible moral agents, and attempt to foster in them a certain, attitudinal response, i.e. moral recognition of the wrongness of their criminal acts. It is through moral recognition that wrongdoers are able to make amends, since moral recognition in combination with shouldering certain burdens is what imbues the wrongdoer's apology with meaning. Return for a moment to the Drunk Driver case in Chapter 4. Brian recognizes that his behaviour, and especially what he did to Jody, was morally wrong. Thus, the burdens he agrees to undertake for the sake of making amends are meaningful to Jody and other members of the community. Brian’s apology is rendered meaningful by way of his moral recognition of the gravity of his offence, combined with whatever burdens he undertakes in order to “right the wrong.”
Before moving on, it is worth briefly situating the restorative justice account in relation to other retributivist themes previously discussed.

5.6.1 – Restorative justice and forfeited rights

Recall that forfeited rights retributivists argue that wrongdoers deserve to suffer punishment for the sake of removing the unfair advantage that accrues to wrongdoers via their non-compliance. Recall that Morris argues that a system of criminal law ought to be “…one in which the rules establish a mutuality of benefit and burden and in which the benefits of noninterference are conditional upon the assumption of burdens.”\(^{218}\) Despite the various controversies surrounding this sort of view,\(^{219}\) it retains intuitive appeal, especially regarding the idea that upholding societal fairness is an aspect of what justifies institutions of criminal justice.

We are now in a position to situate the restorative justice approach in relation to the forfeited rights account; restorative justice is, like the forfeited rights account, about wrongdoers making amends. It is unfair if wrongdoers are allowed to proceed to live their lives as if they had never wronged others. If wrongdoers are unwilling to live up to what they owe to their victims and the community, then the state ought to intervene in their lives for the sake of persuading them to make amends. The forfeited-rights account is a way of capturing a similar sort of intuition, namely, that it is unfair if wrongdoers fail to recognize their wrongdoing as wrongdoing, and fail to make amends with those

\(^{218}\) See Morris, “Persons and Punishment,” p.42.
\(^{219}\) See Chapter 3, section 3.
whom they have wronged. The idea of stripping the wrongdoer of her supposed “unfair-advantage” might be construed as a way of capturing the intuition that the state ought to intervene in the lives of wrongdoers in order to persuade them to make amends, toward the further end of repairing the harm inflicted upon the community by virtue of criminal wrongdoing.

5.6.2 – Restorative justice and punitive emotions

Recall that punitive emotions retributivists argue that criminal punishment is justified in part as an expression of the “hatred and … vengeance” that criminal wrongdoing “excites in healthily constituted minds.” Along similar lines, Moore argues that moral guilt is the appropriate response to crime on the part of the offender. Therefore, he argues, if you or I were guilty of a crime, the extreme extent of our own guilt would be such that we would undoubtedly endorse a punitive response to our own wrongdoing, i.e. we would rightly judge that we ourselves deserve to suffer. Thus, Moore argues, if we are to treat wrongdoers as rational and responsible moral agents such as ourselves, then we should also endorse a punitive response to criminal wrongdoing in the general case.

In my view, Moore is correct in claiming that wrongdoers ought to feel guilty. According to the restorative justice account, wrongdoers have an obligation to recognize their wrongdoing as such, i.e. as wrongdoing, and hence they ought to feel appropriately guilty for having wronged others. However,

---

221 Moore, *Placing Blame*, p.149.
according to the restorative justice account, the wrongdoer’s feelings of guilt ought to motivate her to action, i.e. to make amends for her wrongdoing by shouldering certain burdens for the sake of reconciling herself with those whom she has wronged. The imposition of suffering alone, by contrast, is in an important sense a passive response to wrongdoing, in that it is a response that can be imposed upon wrongdoers without appeal to their moral agency. In my view, therefore, merely suffering the harm of criminal punishment is an inadequate response to wrongdoing on the part of the wrongdoer. On its own, suffering does nothing to repair the damage that crime does to the community. However, suffering guilt in combination with shouldering certain burdens directed at making amends, can imbue an apology with meaning, i.e. suffering guilt combined with the burden of apology might be valuable for the sake of repairing the wrongdoer’s relationship with her community. A restorative justice approach, therefore, views the imposition of suffering by an institution of criminal punishment as valueless in itself. If, however, the imposition of suffering brings about remorse on the part of the wrongdoer, combined with an effort to make amends, then the imposition of suffering (as well as the suffering of guilt) are valuable as a means to the ends of restorative justice.

It follows that if suffering guilt renders the wrongdoer completely inert (for example due to an extreme experience of self-loathing\textsuperscript{222}) then suffering guilt can

\textsuperscript{222} Of course, such an experience might be the result of a restorative justice process that aims at bringing the wrongdoer to recognize the moral gravity of her wrongdoing. However, a restorative justice approach would prescribe action on the part of the wrongdoer (in the form of making amends) rather than suffering guilt as a valuable end in itself.
actually become counterproductive to the achievement of the ends of restorative justice; the wrongdoer might be so guilt-stricken that she refrains from any attempt at reconciling herself with the community, or worse, she might come to view her life as completely devoid of value and no longer worth living. The value of suffering guilt derives from its role in achieving further valuable ends, i.e. the ends of restorative justice.

Along similar lines, the value of the community expressing outrage given some instance of criminal wrongdoing, as Stephen suggests is appropriate, might likewise be counterproductive to the ends of restorative justice. While holding wrongdoers accountable to the demands of morality is the appropriate response to crime on the part of the community, when the expression of blame rises to the level of unbridled outrage and hatred on the part of the community toward the wrongdoer, it can impede the achievement of the ends of restorative justice. In such a case, the wrongdoer may become resentful of the community who expresses hatred towards her. She may come to view the community as her enemy, rather than as a community within which her continued membership is valued.

5.6.3 – Restorative justice and deserved censure

Recall that Nozick argues that “The wrongdoer has become disconnected from correct values, and the purpose of punishment is to (re)connect him. It is not that this connection is a desired further effect of punishment: the act of retributive
punishment itself effects this connection. Expressive retributivists such as Nozick, von Hirsch and others, take issue with the penal aim of attempting to induce moral recognition in wrongdoers. For expressive retributivists, the act of retributive punishment itself reconnects the wrongdoer with correct values, albeit “...in a second best way.” According to the restorative justice account, by contrast, it is the wrongdoer who has a duty to recognize the moral reasons not to offend, and hence also the wrongness of her criminal acts. Therefore, the restorative justice account is much more ambitious in its aims concerning how criminal justice institutions ought to relate to wrongdoers.

Nevertheless, expressive retributivists are correct in claiming that the community has an obligation to hold wrongdoers accountable, in the sense of censuring their criminal misconduct. However, we should resist jumping to the conclusion that retributive punishment is always the most appropriate mode of expressing the condemnation of the community. A restorative justice account views the permissibility of the imposition of hard treatment (as the particular mode of expressing blame) in a way that is tied to the secondary duties of offenders. Only insofar as the wrongdoer is liable to be harmed for the sake of convincing her to fulfill her obligations to others may she be subjected to public censure in the form of penal hard treatment. And even then, as argued in Chapter 4, in any given case, other considerations, i.e. considerations independent of the culpability of the wrongdoer, ought to factor into our all things

---

considered judgement concerning the appropriate mode of expressing our collective judgement of moral condemnation.

However, expressive retributivists are likely to complain that the restorative justice account is much too ambitious. Many expressive retributivists deny, for example, that it is the appropriate role for the state to meddle in the moral constitution of its citizens. For this reason, they argue that the expression of blame goes far enough; we should not, they argue, through penal policy and the imposition of hard treatment, attempt to actually bring wrongdoers to recognize that they are guilty of moral wrongdoing. I respond to this objection in the next Chapter.

5.7 – Conclusion

In this Chapter, I defended a non-retributivist, restorative justice theory of punishment grounded in the duties of offenders. The value of an institution of criminal punishment, in my view, derives from the more basic valuable ends of restorative justice. The duties of offenders, I argued, generate limited permissions for the state to impose hard treatment upon wrongdoers for the sake of persuading them to live up to what they owe to others in both the attitudinal and expressive senses.

Subsequently, I outlined some limitations associated with my view. We should not view the restorative justice account as prescribing the imposition of hard treatment in all cases. Rather, we should consider the use of alternatives, if and when such alternatives are deemed to be more effective in convincing
wrongdoers to live up to what they owe to others in the requisite sense. Along similar lines, it is likely that the forms of punishment rendered permissible by the restorative justice account, given the restrictions set out by the secondary duties of offenders, will be extremely limited in scope. Any form of criminal punishment that fails to include genuine, substantive opportunities for wrongdoers to discharge their secondary obligations of moral recognition cannot be justified. Finally, it is unlikely that any existing system of criminal punishment can be said to live up to the ideal set out by my account. Permissible criminal punishment for the sake of restorative justice would be challenging to implement. Given the nuanced and complex character of the ethics of criminal punishment, this should come as no surprise.

Finally, I attempted to situate my account in the context of some well-known retributivist themes. The restorative justice account and forfeited-rights retributivism have in common the intuitively appealing demand that criminal justice institutions uphold fairness in society. Also, both accounts emphasize in their own way the idea that the appropriate response to wrongdoing on the part of the community is to demand that wrongdoers make amends.

The restorative justice account and punitive emotions retributivism have in common the idea that wrongdoers ought recognize their wrongdoing for what it is, i.e. wrongdoing, and to feel appropriately guilty for wronging others. The two accounts differ in how they view the value of the suffering of offenders. For Moore and others, the suffering of wrongdoers is what is valuable about criminal punishment, while the restorative justice account demands more from the
suffering of wrongdoers; the suffering of wrongdoers is only valuable, in my view, insofar as it motivates the process of moral reconciliation between wrongdoers and their communities. As I argued in this Chapter, this end is not something that can be brought about by sheer force. The ends of restorative justice must ultimately flow from the agency of the wrongdoer.

The restorative justice account and expressive versions of retributivism have in common the notion that the community has an obligation to hold wrongdoers accountable. However, the restorative justice account insists that if the imposition of hard treatment is to be rendered permissible, it must be tied to an effort to convince wrongdoers to live up to what they owe to others. Expressive retributivists are likely to complain that such an aim is too ambitious, i.e. that it is beyond the appropriate purview of the state to attempt, through the use of force, to induce moral recognition and repentance in wrongdoers. In the next Chapter, I turn to a series of objections in the literature surrounding the penal aim of attempting to foster moral recognition in wrongdoers.
Chapter 6 – Defending punishment for the sake of moral recognition

6.1 – Introduction

In the previous Chapter, I defended a restorative justice theory of punishment grounded in the duties of offenders. The restorative justice approach focuses on healing or repairing the damage done to the community resulting from criminal forms of wrongdoing. I argued that the secondary duties of offenders render permissible, but also ought to constrain, our use of hard treatment for the sake of pursuing the ends of restorative justice. In my view, criminal punishment, understood as a mode of forceful, public censure, should aim to persuade wrongdoers to live up to what they owe to others in the attitudinal and expressive senses. I argued that the forms of punishment we impose upon wrongdoers should address them as rational and responsible moral agents, and attempt to foster in them a certain, attitudinal response, i.e. the response of moral recognition. It is through a process of moral recognition and apology that wrongdoers are able to make amends; moral recognition, combined with shouldering certain burdens, is what imbues the wrongdoer's apology to the community with meaning.

In my view, wrongdoers incur second-order obligations (1) to recognize the moral status of their victims and others, and hence the wrongness of their crimes, and (2) to express or communicate their recognition to the community via

\[225\] See Chapter 4, section 2.
meaningful apology. Such obligations, we noted, generate limited permissions for a third party, e.g. the state, on behalf of the normative community, to interfere in the lives of wrongdoers in order to persuade and enable them to live up to what they owe to others. Therefore, in my view, in order to be permissible, a system of criminal punishment must aim at what I call the ends of restorative justice.226

According to the restorative justice account, the purpose of an institution of criminal punishment, if and when such an institution ought to be considered permissible, is to forcefully communicate blame or censure toward wrongdoers, toward the further ends of restorative justice. I have argued that punishment is not constitutive of the ends of restorative justice, but when considered against possible alternatives, an institution of criminal punishment may be valuable for the sake of pursuing such ends. The account I defend is therefore non-retributive in character, in that my account finds no necessary relationship between criminal punishment and the ends of restorative justice.227 Rather, in my view, the ends of restorative justice are tied to an account of the secondary obligations of offenders. However, according to a restorative justice approach grounded in the duties of offenders, an institution of criminal punishment is prescribed as one institutional possibility among others that is valuable for the sake of pursuing the ends of restorative justice. Therefore, it is worth responding to several well-known objections to the view that criminal punishment ought to aim at fostering

---

226 As previously discussed, I refer to the aims of (1) fostering moral recognition in wrongdoers, and (2) fostering reparations between wrongdoers and their victims as co-members of a shared normative community collectively as the ends of restorative justice. See Chapters 4 and 5.

227 For this argument, see Chapter 4, section 5.
recognition in wrongdoers, and reparations between wrongdoers and their victims as co-members of a shared normative community.

In this Chapter, I defend the restorative justice account against some prominent objections. In particular, I respond to several objections to the view that an institution of criminal punishment ought to aim at fostering moral recognition in wrongdoers. A restorative justice approach to the problem of punishment is famous in the contemporary philosophy of punishment literature. However, philosophers have recently advanced several well-known objections to the view that an institution of criminal punishment ought to aim at fostering moral recognition in wrongdoers. It is the purpose of this Chapter, therefore, to defend the restorative justice theory of punishment against such objections. In what follows, I address four separate worries: that punishment for the sake of fostering moral recognition in wrongdoers fails to provide a justification for punishing already repentant and predictably defiant offenders (section 3), that the restorative justice account fails to make room for a principle of proportionality (section 4), that focusing on moral recognition as opposed to crime prevention in the context of criminal justice fails to provide sufficient protection against wrongdoing, and therefore puts society at unnecessary risk (section 5), and finally, that it is not the appropriate role for the state to aim to induce moral recognition in wrongdoers via an institution of criminal punishment (section 6).

As argued in Chapter 5, I share with Duff a commitment to the view that justifiable criminal punishment ought to aim at fostering moral recognition in wrongdoers. For a discussion of how my view differs from Duff’s account, see Chapter 4, sections 4-5.
6.2 – Punishing the already repentant and the predictably defiant

Two related objections take issue with the penal aim of fostering moral recognition in wrongdoers as such: (I) one might argue that the restorative justice account fails to provide a compelling justification for punishing those who are already remorseful and/or repentant for their wrongful conduct. (II) Alternatively, one might argue that the restorative justice account fails to provide a compelling justification for punishing those who are incorrigible, or predictably and unfailingly defiant, in the sense of being uninterested in the moral reasons not to offend.

Von Hirsch sums up these two complaints nicely:

Were inducing penitent reflection the chief aim [of an institution of criminal punishment], as R.A. Duff has claimed, there would be no point in censuring actors who are either repentant or defiant. The repentant actor understands and regrets his wrongdoing already; the defiant will not accept the judgement of disapproval which the censure expresses. Yet we would not wish to exempt from blame either the repentant or the seemingly incorrigible actor.\(^{229}\)

Therefore, for von Hirsch and others, if fostering moral recognition in wrongdoers is the chief justificatory aim of a system of criminal punishment, then there would be "no point" in punishing those who are either already repentant or predictably defiant.

6.2.1 – Punishing the already repentant

First consider the case of the already repentant offender. In my view, von Hirsch is correct that we would not want to exempt those who are already repentant from being held accountable for their wrongful conduct. To see why we

\(^{229}\) See von Hirsch, Censure and Sanctions, p.10.
should not accept the conclusion that punishment in such cases would have “no point,” consider the following hypothetical case:

Already Repentant: Bob is guilty of violently assaulting Sally. More than a full year later, Bob is arrested for his crime. During the intermediate year, Bob experiences significant hardships in his life (suppose that one of Bob’s family members is victimized in a similar way) making it abundantly clear to Bob the wrongfulness of his conduct toward Sally more than a year ago. Bob now deeply regrets that he assaulted Sally.

Now, consider the attitude Bob might adopt toward his own punishment for assaulting Sally. One possibility, as von Hirsch suggests, would be for Bob to view his own punishment as having “no point.” To elucidate this idea, we might suppose that Bob is brought to trial and his defense takes something like the following form: “Yes, I committed the crime in question, but I am truly remorseful. Due to my experiences in the past year, I have come to fully understand and recognize that what I did to Sally more than a year ago was horribly wrong. I am filled with remorse, and I am willing to beg for forgiveness from Sally, her family and friends, and the moral community more generally.” Bob then goes on to conclude, “Punishing me, therefore, would be without moral justification or purpose.”

It should be apparent, I hope, why Bob’s argument is flawed. His conclusion, namely, that his punishment would be “without purpose” calls into question the authenticity of his previous claim regarding his recognition (and his willingness to communicate his recognition) of the wrongfulness of his conduct. In my view, the clearest way of responding to von Hirsch’s argument is as follows: von Hirsch fails to account for Bob’s secondary obligation of moral recognition in the attitudinal and expressive senses. While the already repentant
wrongdoer, by definition, has already fulfilled his secondary obligation of moral recognition to others in one, quite important sense, i.e. in the attitudinal sense, as argued in Chapter 5, moral reform on the part of offenders is arguably not enough for them to fully live up to what they owe to others. Wrongdoers arguably also owe it to others to meaningfully demonstrate their recognition of the moral gravity of their crimes in the form of a meaningful apology. If the already repentant wrongdoer were to view his own punishment, as Bob does, as having “no point,” then he would be failing in an important sense to live up to what he owes to his victims and others, i.e., he would be failing in his secondary obligation of moral recognition in the expressive sense.

However, one might push the objection even further: what of those offenders who are not only already remorseful, but have also already suffered a form of burden in their private life, enabling their expression of a meaningful apology? Recall that Bob has suffered hardships in his life that trigger his recognition of the wrongness of his criminal acts. If Bob verbally expresses an apology, coupled with reference to the hardships he has suffered in his private

---

230 See Chapter 5, section 3.
231 Importantly, as argued in Chapters 4 and 5, suffering punishment might not always be the most appropriate mode of expressing blame in the case of already repentant offenders. Suffering punishment is also not the only way for the wrongdoer to shoulder burdens for sake of expressing meaningful apology (cf. the Drunk Driver case in Chapter 4). So, in other words, if the already repentant offender were to view the burden of expressing meaningful apology (however the burden is undertaken, i.e. whether by suffering punishment or by suffering some other form of burden aimed at making amends) as having “no point,” he would be, in an important sense, failing to live up to what he owes to others, i.e., by failing in his obligation of moral recognition in the expressive sense. See Chapter 4, section 5.
life, then should the community not accept his apology in such a case as sufficient?

The problem with this sort of argument is that criminal forms of wrongdoing damage the relationship between the wrongdoer and the normative community. Therefore, a form of burden that addresses the community is required in order to enable an appropriate apology that aims at repairing the impaired relationship between the wrongdoer and her community. The burden of meaningful apology must be community-oriented: suffering the burden of apology must be directed toward the community and coupled with explicit recognition of the wrongness of one's criminal acts. Suffering burdens in one's private life, therefore, are not the sort of burdens that can enable reparations between wrongdoers and their communities. The wrongdoer must undertake some form of burden that aims to reconcile her relationship with those whom she has wronged, either by suffering the burden of communicative punishment, or by suffering some other burden for the sake of making amends, i.e. a burden directed toward repairing the relationship between the wrongdoer and those whom she has wronged. Discharging one’s secondary obligation of moral recognition in the expressive sense, therefore, involves a meaningful public apology on the part of the wrongdoer. Such an apology must be rendered meaningful by undertaking a form of burden that addresses the community in some way, rather than an apology that derives its meaning from hardships in one’s private life.\footnote{232}{Hardships in one’s private life can, however, influence moral reform on the part of offenders, as in the case of Bob in \textit{Already Repentant}. As such, private hardships should not be considered irrelevant to the relationship between...}
6.2.2 – Punishing the predictably defiant

Now consider the case of the predictably defiant offender. Again, von Hirsch is correct that we would not want to exempt those who are predictably defiant from being held responsible for their criminal acts. But, according to the restorative justice account, why should we punish individuals who we know are extremely unlikely to be moved by moral reasons? Known members of organized crime, for example, are often thought of as falling within this sort of category. Perhaps in such cases, one might argue, incapacitation and specific deterrence are the only appropriate goals to be pursued by an institution of criminal punishment.

The problem with this sort of argument, in my view, is the empirical dimension of the claim that certain wrongdoers are, or ought to be considered, incorrigible. It seems to me that empirical reasons for considering certain wrongdoers as entirely beyond the pale, in the sense of being irredeemably evil, are rarely going to be strong enough to establish with certainty the conclusion that we as a community ought to give up on such individuals with respect to the future possibility of moral reform.233

Furthermore, the argument supplied by von Hirsch is vulnerable to the following response: the incorrigible offender, like other offenders, has an

---

233 See Duff, “Penal Communications,” p.54.
obligation to recognize and express recognition of the moral status of her victim(s) and others. The fact that a certain wrongdoer is predictably unlikely to be moved by moral reasons, and so unlikely to express genuine recognition and remorse, does not relieve her of her responsibilities as a wrongdoer. Therefore, when interference for the sake of fostering recognition is unlikely to be effective, it nevertheless ought to be considered permissible in a wide variety of cases, since the duties of offenders create limited permissions for others to intervene in the lives of wrongdoers for the sake of attempting to persuade them to recognize the moral status of others, and hence the wrongness of their crimes.\footnote{Cf. notes 208 and 213 above. Also, as will be reiterated in what follows, if punishment is unlikely to have the requisite persuasive effect on the wrongdoer, then alternatives to punishment (alternative forms of coercive interference) should be considered alongside (or in place of) punishment, for the sake of pursuing the ends of restorative justice.}

As argued in Chapters 4 and 5, if criminal punishment as a mode of intervention is unlikely to be effective, we should look to other modes of holding wrongdoers accountable, e.g. a process of shared moral inquiry,\footnote{See Chapter 4, section 5. This is not to claim, however, that we should let defiant offenders “off the hook,” so to speak. As will be discussed in what follows, while we might adjust the mode of expressing the judgement that a certain wrongdoer is blameworthy, and hence ought to be held accountable to the community, the severity of the burdens we impose upon her should be proportionate to the onerousness of her secondary obligation to make amends. Therefore, defiant wrongdoers should not necessarily be treated with leniency when compared with other offenders guilty of similar offences. This point will be further discussed in what follows.} either in addition to or alongside punishment, for the sake of efficacy with respect to the aim of fostering moral recognition. It is important to note, however, that the unwillingness of certain offenders ultimately to be convinced, i.e. to recognize the moral status of others despite forceful modes of persuasion, must be respected.
As Duff rightly points out, respecting the agency of wrongdoers requires that we allow incorrigible offenders to remain, after all, unpersuaded. However, in my view, such wrongdoers nevertheless ought to be considered liable to be harmed for the sake of attempting to persuade them in the requisite sense.

6.2.3 – Punishment and the problem of psychopathic offenders

As argued above, we shouldn’t give up hope when confronted with certain offenders who are predictably defiant, in the sense of being resistant to the force of moral reasons. Our interventions should, I argued, aim at moral reform, even if criminal punishment is not the chosen mode of holding incorrigible wrongdoers accountable, in the sense of attempting, through the use of force, to persuade them to see the force of moral reasons. However, we should consider an additional problematic case, namely, offenders who are not simply predictably unlikely to see the force of moral reasons not to offend, but rather psychologically unable to see the force of such reasons.

236 Cf. Chapter 5, section 4. This delicate balance, i.e. between interference for the sake of moral recognition and respect for the agency of wrongdoers, is essential to the aim of pursuing the ends of restorative justice in a way that is compatible with respecting the status of wrongdoers as rational and responsible moral agents. See for example Duff, “Penal Communications,” p.54. For Duff, "...punishment must aim at rational moral persuasion, not coerced (and hence inauthentic) acceptance: thus it must always address the offender as a rational moral agent..." Furthermore, the duties of offenders do not prescribe indeterminate modes of criminal treatment for the sake of pursuing the ends of restorative justice. This point will be discussed in the following section.
The problem of psychopaths is well known in the contemporary philosophical literature on moral responsibility and blame. On the one hand, psychopaths are able to grievously wrong others, and hold attitudes that are morally reprehensible. On the other hand, however, since such individuals are (by definition) unable to feel the force of moral reasons not to offend, they seem to lack an important aspect of moral agency that is, at least at first glance, necessary for moral accountability. As Watson writes,

The phenomenon of psychopathy brings to the fore a distinction between ... two “faces” of responsibility. We rightly predicate viciousness of the attitudes and conduct of psychopaths (“dishonest,” “abusive,” “manipulative,” “hostile,” “mean”), and we naturally respond to them accordingly. This is the attributability face. And yet they lack the capacity for moral reciprocity or mutual recognition that is necessary for intelligibly holding someone accountable to basic moral demands and expectations. This capacity delineates the boundaries of moral accountability, and thereby the other face of responsibility.

How, one might ask, is the restorative justice account able to accommodate psychopathic offenders? The problem, it seems to me, is as follows: since ought reasonably implies can, it is reasonable to assume the sort of secondary obligations specified by the restorative justice account do not apply in the case of psychopaths. If psychopaths are psychologically unable (i.e. in the sense of lacking a capacity) to partake in moral reciprocity, and thus to feel the force of moral reasons not to offend, then it is not reasonable to assume that they

---

238 I.e. psychopaths hold attitudes that fail to respect the basic moral status of others.
239 I.e. in the sense being held at fault and accountable for their actions toward others.
have an obligation to do something that is not by definition in their power to do, i.e. to recognize the moral status of others, and hence the moral wrongness of their crimes.\textsuperscript{241}

Therefore, with such considerations in mind, we should view the justification for interference in the case of psychopaths as reverting back to the importance of their primary obligations toward others, i.e. their obligations not to (re)offend. Psychopathic offenders have obligations, like the rest of us, to refrain from violating or disregarding the basic moral status of others. Lacking the capacity to feel the force of moral reasons not to violate or disregard the basic interests of others does not imply that violating or disregarding the basic interests of others in such cases is permissible. On the contrary, in such cases, once we have established that (1) an offender is a psychopath, and therefore that (2) such an offender would have secondary obligations of moral recognition if otherwise psychologically disposed, we should view the purpose of coercive interference as reverting back to a focus on the wrongdoer’s first-order obligations, i.e. to refrain from (re)offending. In such cases, therefore, the aim of coercive interference should be that of \textit{specific deterrence}, i.e. the aim of providing psychopathic offenders with prudential reasons not to reoffend.\textsuperscript{242}

\textsuperscript{241} And if the obligation of moral recognition in the attitudinal sense does not apply, then since discharging the attitudinal dimension of the obligation is a condition for discharging the expressive dimension, the expressive dimension of the obligation likewise does not apply.

\textsuperscript{242} Once again, the distinction between specific and general deterrence is relevant, here: specific deterrence refers to providing the wrongdoer with prudential reasons not to offend, while general deterrence refers to providing the general public with prudential reasons not to offend, i.e. for fear of the harm or burden of criminal punishment. See Chapter 2, note 9. As will be discussed in
But, one might ask, what about moral accountability? Should psychopathic offenders, if unable to feel the force of moral reasons, nevertheless be subjected to punishment, i.e. a mode of treatment that aims at holding the wrongdoer accountable via forceful, public censure? Perhaps somewhat counter-intuitively, in my view, the correct answer is “Yes.”

First, there are good reasons to think that, despite the psychological fact of being unable to feel the force of moral reasons not to offend, the nature of the agency involved in psychopathic wrongdoing nevertheless makes it appropriate to hold psychopathic wrongdoers accountable for their wrongful actions toward others. Psychopathic offenders, despite their diminished form of moral agency, nevertheless possess the ability to assess reasons, and hence may be permissibly held accountable to others for their wrongful actions. Furthermore, in such cases, we have good reasons to make use of criminal punishment as a mode of expressing blame, since, while such offenders are unable to view public censure as expressing moral reasons not to offend, both the force of censure and the hard treatment that are constitutive of communicative, criminal punishment are likely to be viewed by such individuals as providing prudential reasons not to reoffend, toward the end of specific deterrence. Furthermore,

what follows, in my view, the special case of psychopathic wrongdoers does not justify pursuing the penal aim of general deterrence, since the vast majority of people are in fact able to feel the force of moral reasons not to offend.

As Scanlon argues, moral criticism does not rely on the wrongdoer’s comprehension of “...moral reasons in particular.” He writes, “A person who is unable to see why the fact that his action would injure me should count against it still holds that this [i.e. the fact that his action would injure me] doesn’t count against it” (Scanlon’s italics). See Scanlon, T.M., *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998) pp.288-9. For a critical discussion of Scanlon’s view, see Watson, “The Trouble with Psychopaths.”
censuring the wrongdoer in cases where we have good reasons to think we are dealing with a psychopath leaves open the opportunity for such an individual to respond to the moral reasons not to offend, in the unfortunate scenario wherein we mistake a psychopathic offender for a merely defiant or incorrigible offender, i.e. such a form of holding wrongdoers accountable is consistent with the secondary obligations of non-psychopathic offenders.

Importantly, on this view, the pursuit of specific deterrence in the special case of psychopathic wrongdoers ought to be considered grounded in the specific first-order obligations of such wrongdoers to refrain from (re)offending. Only in the special case of wrongdoing, and only when we are sure that we are dealing with a psychopathic wrongdoer, should we view the goal of coercive interference as oriented towards specific deterrence. As argued in Chapter 5, typically, interference for the sake of compelling individuals to live up to their first-order obligations to refrain from wronging others is not, at least in all possible cases, desirable; it would involve too much state interference in our personal lives. However, in the special case of wrongdoing, and in a subset of such cases wherein the wrongdoer is determined to be a psychopathic offender, we should view interference as permissible for the same reasons that interference ought to be considered permissible in cases wherein the second-order obligations of moral recognition apply – to compel wrongdoers to live up to what they owe to others as wrongdoers. The difference in the case of psychopathic offenders, since they lack a certain capacity for feeling the force of moral reasons, is that their second-order obligations (their enforceable obligations generated by
wronging others) revert back to the first-order requirement that they refrain from wronging others.\textsuperscript{244}

6.3 – Proportionality, moral reform, and the duties of offenders

As argued in Chapter 4, there is no necessary relationship between a system of criminal punishment and the ends of restorative justice. It is conceptually possible that wrongdoers might fulfill their secondary obligations of moral recognition in ways that do not involve being subjected to the harms of a system of criminal punishment.\textsuperscript{245} Even if wrongdoers incur certain duties, we should resist jumping to the conclusion that a system of criminal punishment is, in principle, the best or only way for wrongdoers to discharge their secondary obligations. It might turn out that a system of criminal punishment is indeed the most efficient means for the sake of pursuing such ends. However, it also might turn out otherwise, i.e., that other forms of coercive treatment, e.g. a compulsory process of shared moral inquiry, are more conducive to fostering recognition in

\textsuperscript{244} In my view, making room for the pursuit of specific deterrence grounded in the duties of offenders in the special case of psychopathic offenders is a better way of capturing von Hirsch’s suggestion that the hard treatment involved in the criminal sanction ought to provide a prudential supplement to the moral voice of the criminal law. See Chapter 3 section 4. The difference is that prudential supplements ought to be considered meaningful reasons to obey the criminal law only in the case of those who are psychologically unable to grasp the force of moral reasons not to offend, e.g. in the case of psychopathic offenders. Otherwise, prudential supplements fail to treat wrongdoers and the general public with the respect they are owed as rational and responsible moral agents.

\textsuperscript{245} This is an implication of the structural features of my account, i.e. that the value of an institution of criminal \textit{punishment}, as I understand it, ought to be considered valuable only as one institutional possibility among others for the sake of pursuing the ends of restorative justice. For a detailed discussion of this feature of my view, see Chapter 4 section 5.
wrongdoers, and so more conducive to allowing for wrongdoers, as wrongdoers, to discharge their relevant secondary obligations.

However, as discussed in Chapter 2, the view that criminal wrongdoers ought to be subjected to reformative criminal treatment as opposed to criminal punishment per se has fallen from grace in the contemporary moral and political philosophical literature.\(^{246}\) Recall that Moore argues,

\[\text{[R]ecasting} \ldots \text{punishment in terms of ‘treatment’ for the good of the criminal makes possible a kind of moral blindness that is dangerous in itself} \ldots \text{[A]dopting such a ‘humanitarian’ conceptualization of punishment makes it easy to inflict treatments and sentences that need bear no relation to the desert of the offender.}\(^{247}\]

Critics of reformative alternatives to criminal punishment point out that justifiable coercive treatment in the case of criminal wrongdoing must fit the crime. A principle of proportionality, such critics argue, is a basic requirement of any defensible theory of criminal punishment, or corrective theory of criminal treatment. It is morally unacceptable, therefore, if a theory prescribes an unduly harsh duration or character of criminal treatment for an offender culpably responsible for a relatively minor offence (say, petty theft). Conversely, it is morally unacceptable if a theory prescribes an unduly minimal or lenient duration or character of criminal treatment for an offender culpably responsible for a...

\(^{246}\)Although this distinction is not central to my inquiry in what follows, it is possible to distinguish between rehabilitative and reformative treatment programs. While reformative treatment aims at reforming (morally, socially, psychologically, etc.) offenders, rehabilitative treatment aims at providing offenders with capacities and opportunities (employment opportunities, educational opportunities, etc.) meant to smooth their transition back into society, or, more ambitiously, to improve their overall level of wellbeing or flourishing. For this point see Duff, *Punishment*, p.5.

\(^{247}\)See Moore, *Placing Blame*, p.87.
relatively egregious offence (say, murder). The burden of criminal punishment, or corrective criminal treatment, therefore, ought to be proportionate to the seriousness of the wrongdoing in question. However, critics argue that defenders of reformatory alternatives to criminal punishment provide no basis for adopting a principle of proportionality.

On the restorative conception of punishment, however, the duties of offenders provide a concrete framework for grounding a principle of proportionality: the weightier the seriousness of the wrongdoing in question, the weightier we should consider the relevant secondary obligations of offenders.

While the attitudinal obligations of wrongdoers seem, at least at first glance, similar across cases, i.e. to reform oneself, and to adopt attitudes (i.e. beliefs, intentions, desires, etc.) that do not involve blatant or passive disregard for the moral status of others, intuitively, the communicative obligations of offenders admit of degrees proportionate to the wrongdoing in question. For example, if I steal your car, one thing that I owe you is a rather weighty apology. If, on the other hand, I murder a member of your family, the apology that I owe you ought to be considered to a much greater degree weighty and onerous. Indeed, in such a case, we might think that I owe you a lifetime of demonstrated remorse and apologetic behavior to even begin to live up to my obligations to you after the fact. Thus, a principle of proportionality ought to be considered built in to any reasonable account of the secondary obligations of offenders. Criminal justice institutions, therefore, ought to reflect our commitment to such a principle, both in the context of criminal punishment and in the context of corrective
alternatives to criminal punishment, e.g. a process of shared moral inquiry. Thus, keeping in mind that human institutions in many cases are only able to roughly or imperfectly bring about the outcomes that we deem consistent with our moral commitments, the duties of offenders prescribe the use of more limited coercive interference in cases involving relatively minor offences, and the use of more significant coercive interference in cases involving relatively egregious offences.

It is worth highlighting at this point a potential further objection regarding corrective alternatives to criminal punishment and proportionality: corrective or educational modes of criminal treatment, one might argue, are much more conducive to enabling the discharge of wrongdoers’ secondary obligations in the attitudinal sense, as opposed to the communicative or expressive sense. One might argue that a form of corrective criminal treatment is much more conducive to convincing wrongdoers to adopt attitudes and beliefs that do not involve moral disregard for others, and much less conducive to enabling wrongdoers to meaningfully apologize to the community and those whom they’ve wronged. But if proportionality is important only in the communicative context, then the worry regarding disproportionate forms of criminal treatment in the attitudinal context remains: we might worry that, on this view, wrongdoers are considered liable to indefinite and indeterminate modes of coercive treatment for the sake of convincing them to reform themselves, i.e. convincing them to fulfill their obligations to others in the attitudinal sense.

However, we should resist this characterization of the implications of grounding the permissibility of corrective alternatives to criminal punishment in
the duties of offenders: while the attitudinal obligation seems to be similar across cases, i.e. to reform oneself and to adopt certain attitudes and beliefs and eschew others, the liability of wrongdoers to be harmed for the sake of achieving such ends is not. A principle of proportionality ought to rein in our use of coercion in both the attitudinal and communicative contexts. In the communicative context, as previously discussed, we ought to consider the burden of apology as proportionate to the seriousness of the offence. Therefore, we ought to consider the liability of wrongdoers to be harmed for the sake of convincing them to discharge their obligations in the communicative sense as likewise proportionate. In the attitudinal context, subjecting wrongdoers to indefinite modes of coercive treatment for the sake of convincing them to reform themselves would fail to treat wrongdoers with the respect they are owed as rational and responsible moral agents. Such forms of treatment would be counterproductive, since we are unlikely to be successful in persuading wrongdoers of the moral reasons not to offend if we are unwilling to be definite with respect to the terms and limitations of their liability to be harmed for the sake of inculcating such reasons. Furthermore, such forms of treatment would fail to respect the condition that moral recognition must ultimately flow from the agency of the wrongdoer. Indefinite modes of corrective, criminal treatment for the sake of fostering recognition in wrongdoers would be too coercive. Coercive intervention ought to attempt to persuade wrongdoers to reform, rather than attempt to extract moral recognition from wrongdoers by sheer force. As such, indefinite modes of corrective, criminal treatment for the sake of fostering moral recognition in wrongdoers would miss
the point, i.e. that wrongdoers as moral agents ought to fulfill their obligations to others in the community.\(^{248}\)

However, a brief caveat is in order: since the account I defend is non-retributivist in character, as argued in Chapter 4, we should distinguish in the context of criminal justice between the judgement that a wrongdoer is blameworthy, and hence ought to be held accountable to the community, and the appropriate expression of such a judgement, given all the relevant considerations. While the former is an obligatory response to criminal forms of wrongdoing on the part of the community, the latter ought to factor in certain considerations that are independent of the culpability of the wrongdoer, including our standing as a community in relation to certain offenders, the extent to which certain offenders are already willing to discharge their obligations to others, as well as the likely effects of our chosen mode of expression on the wrongdoer, her victim(s), and the community.\(^{249}\) Therefore, in abandoning our reliance on a retributivist theory of moral responsibility and blame, we should be willing to relax the proportionality constraint on permissible harming in the context of criminal justice. As a matter of fairness, on this view, strict proportionality ought to be weighed against the other normatively salient features of our chosen mode of expressing the judgement that some offender is blameworthy, and hence ought to be held accountable to others. Nevertheless, since a proportionality constraint is built in, so to speak, to any reasonable construal of the duties of offenders, strict proportionality ought to be the baseline according to which deviations must

\(^{248}\) Cf. the case of *Manipulation* in Chapter 5, section 4.
\(^{249}\) For this argument see Chapter 4, section 5.
be justified. Furthermore, while we ought to aim for similarity across cases with respect to the severity or onerousness of our interventions, the appropriate mode of intervention might vary across cases, for the sake of pursuing the ends of restorative justice.\textsuperscript{250} This point will be further discussed in the final section (i.e. section 6)

6.4 – Restorative justice and societal protection

In his recent book \textit{The Ends of Harm}, Victor Tadros takes issue with the view that an institution of criminal punishment ought to aim at fostering moral recognition in wrongdoers. For Tadros, general deterrence is the justifying aim that an institution of criminal punishment ought to pursue.\textsuperscript{251} He writes,

The suspicion that we might have about the communicative view is that whilst a great deal of concern is shown to respect defendants and offenders as autonomous agents, insufficient concern is shown for citizens in protecting them from the harms and wrongs that they might be more or less inclined to suffer depending on whether, how, and whom we punish. We have a duty to protect each other from harm and that duty seems to motivate the development and maintenance of criminal justice institutions, but on the communicative view this motivation is suppressed.\textsuperscript{252}

The problem, for Tadros, is that reforming criminal justice institutions along the lines specified by the restorative justice account would likely compromise our security, by failing to provide the general public with prudential reasons not to

\textsuperscript{250} Admittedly, it is difficult, although in my view not impossible, to measure the severity of differing modes of intervention against one another. At the very least, our interventions should be clear about the justifiable upper limit in terms of severity for any given crime in order to prohibit indefinite modes of criminal treatment in any context.

\textsuperscript{251} See Tadros, \textit{The Ends of Harm}, Ch. 12.

\textsuperscript{252} See Tadros, \textit{The Ends of Harm}, p.89.
offend. He argues that such an account over-emphasizes respect for the agency and autonomy of wrongdoers, at the expense of considerations surrounding the security and protection of society at large. On this point, he argues, “…a communicative theory of punishment is a success, but not in the real world.”

In other words, for Tadros, the moral aims of the restorative justice account (i.e. fostering recognition in wrongdoers, and reparations between wrongdoers and their communities) are sound in theory, but too modest to justify a system of criminal punishment in practice, given the tremendous costs associated with administering and maintaining such a system. He writes that the aim of fostering moral recognition in wrongdoers… could be central to the justification of state punishment only in a world that is significantly unlike ours, a world where the costs of punishment are less than they can plausibly be in our world. It is only then that the legitimate but modest goal of recognition could be sought through institutionalized coercion.

The problem for Tadros, therefore, is that the penal aim of fostering recognition in wrongdoers leaves to one side “protecting people from criminal offending.” To elucidate this point, Tadros asks the reader to entertain the following thought experiment:

To see this concern clearly, consider how we would react to a proposal to abolish the criminal justice system. Undoubtedly, one reaction that we would have is that if the proposal were adopted we would then lack a proper public forum for the condemnation of offenders as well as a mechanism through which we could hope that offenders will become reconciled with their communities. But a much more profound reaction would be that if we were to abolish the

---

255 See Tadros, *The Ends of Harm*, p.89.
criminal justice system we would lack adequate protection against offending.\textsuperscript{256}

What might be dubbed ‘the problem of insufficient protection,’ is, at its basis, contingent on empirical considerations. The relationship between reforming any actually existing criminal justice system (e.g. toward the ends specified the restorative justice account) and a rise or fall in crime rates and societal insecurity is an empirical question. It is, however, an empirical question that I think defenders of a restorative justice account should take seriously. Even if we did not go so far as to abolish the criminal justice system, as Tadros suggests we imagine, but sought instead to radically reduce its scope and radically reform its practices, we would then face the objection that in doing so, we might be compromising the security of our society by relinquishing the deterrent aspect of a more traditional approach to criminal punishment.

It is important to note that general deterrence, not societal protection, is what the restorative justice theorist appropriately takes issue with. Aiming at general deterrence via criminal punishment is problematic from the perspective of a restorative justice approach, since harming wrongdoers for the sake of providing people with prudential reasons not to offend militates against the aim of fostering moral recognition in wrongdoers, and thereby providing the general public with morally relevant reasons not to offend. Punishing wrongdoers for the sake of general deterrence presents wrongdoers and others with morally irrelevant reasons not to offend, i.e. to avoid the harms of an institution of

\textsuperscript{256} \textit{Ibid.}
criminal punishment.\textsuperscript{257} We should therefore distinguish between general deterrence as a penal aim, and the broader aim of crime prevention, to which a system of general deterrence is but one possible means. Recall that Duff argues that different theories of punishment focus on different penal aims, such as deterrence, incapacitation, rehabilitation, etc., all for the sake of crime prevention.\textsuperscript{258} Therefore, we should not construe a restorative justice approach as uninterested in societal protection. Rather, such an approach takes issue with the pursuit of societal protection via a system of punishment that aims at general deterrence.

To be clear: setting up a system of punishment that is designed to deter the public from committing crimes is only one possible way of preventing crime by means of criminal justice institutions. Another possible mechanism related to crime prevention, which is more conducive to the penal aim of fostering moral recognition in wrongdoers, is an effort to reduce recidivism by means of fostering moral recognition in wrongdoers. Reducing recidivism rates directly contributes to the security of a society by ensuring that those who offend refrain from reoffending, and furthermore, that they do so for the right reasons. This provides an avenue for allowing wrongdoers as well as the general public to aid in the collective project of societal protection as rational and responsible moral agents capable of comprehending the force of moral reasons.

\textsuperscript{257} For the point about deterrence and relevant reasons not to offend, see Duff, “Penal Communications,” p.15. 
\textsuperscript{258} See Duff, \textit{Punishment}, pp.4-5.
Now return to the problem of insufficient protection. Suppose we were to reform our approach to criminal justice institutions so as to align with the aims specified by the restorative justice account. Furthermore suppose, as Tadros imagines, that we were to subsequently experience a significant increase in crime and societal insecurity. How, one might ask, would an advocate of the restorative justice approach suggest that we deal with such a scenario?

As argued in Chapter 4, our choice as a community concerning the appropriate mode of expressing our collective judgment of blameworthiness in any given case should incorporate certain forward-looking considerations. Among such considerations ought to be our standing as a community in relation to the wrongdoer, and the consideration of whether the wrongdoer is already willing to fulfill her secondary obligations, i.e. whether or not she is already truly remorseful and apologetic. We should also consider, I argued, the likely effects of the mode of expression on the wrongdoer, her victims, and the community more generally. In the context of wrongdoers, we should consider whether or not our chosen mode of expression is likely to be conducive to fostering penitent reflection, i.e. whether it is likely to persuade wrongdoers to reform themselves. In the context of the community, when adopting a particular penal policy, we should consider the likely effects of our chosen policy on the community: a drastic increase in crime would indicate that the community is not responding to the moral message inherent in the criminal law and its prescribed sanctions. Therefore, in light of such an outcome, we should consider adjusting the mode of expressing our collective condemnation of crime in order to better convey the
moral message inherent in the criminal law to both wrongdoers and the community.

Importantly, this is not to claim that, in such a scenario, our penal policies should pursue general deterrence rather than the ends of restorative justice. In order to be consistent with the aim of fostering recognition in wrongdoers, our penal policies must seek to convey to the public morally relevant reasons not to offend. Therefore, when considering alternatives, if it is determined through empirical investigation that a system of communicative criminal punishment is more likely to get the message across to the general public in comparison with some other, non-punitive (coercive) alternative, e.g. a deliberative or educational alternative, then we have a defeasible reason to institutionalize such a system. However, the effectiveness of such a system in preventing crime is, as I have been stressing, an empirical question, as are the effects of our chosen mode of holding wrongdoers accountable with respect to fostering in them moral recognition. We should not assume a priori that any particular institutional approach to the problem of criminal wrongdoing is going to be the most effective at either fostering recognition in wrongdoers, or preventing crime via providing the general public with morally relevant reasons not to offend.

Importantly, it might turn out that a marginal increase in crime ought to be considered a reasonable burden, in that a marginal increase in crime, if it follows from a system that provides the general public with morally relevant reasons not to offend, might be morally preferable to a system that achieves a greater degree of crime prevention yet relies solely on general deterrence, and hence
compliance with the law through sheer fear and unbridled coercion. Furthermore, although admittedly this is a somewhat speculative claim, the notion that crime might radically increase as a result of institutionalizing a restorative justice approach seems unlikely, due to the real-world facts surrounding the overwhelming percentage of people who are, after all, psychologically disposed such that they are able to feel the force of moral reasons not to offend.\textsuperscript{259}

Finally, it is important to point out that if a deterrent approach to criminal justice is deemed necessary in a particular context to avoid high rates of crime and societal insecurity, then we have reasons to hold that such an approach is, from a restorative justice perspective, a necessary evil. While the restorative justice approach provides a morally legitimate framework for criminal punishment, if we are forced in some context by necessity to pursue general deterrence rather than restorative justice, we should view our practice of punishing criminals as, in an important sense, necessary but morally illegitimate. We would have reasons in such a context to pursue by other means the sort of society where the preconditions for a restorative justice approach to criminal justice are met,\textsuperscript{260} so that we are able to abandon the morally questionable practice of pursuing general deterrence via a system of criminal punishment.

It is plausible that other societal factors, including education, healthcare, poverty relief, etc. are related to a societal context in which crime rates are

\textsuperscript{259} See Watson, “The Trouble with Psychopaths,” p.309. In my view, we should not treat the vast majority of people as though they are sociopaths, i.e. potential offenders unable to feel the force of moral reasons not to offend, and unwilling to conduct themselves accordingly. This point will be further discussed in what follows.

\textsuperscript{260} Cf. Duff, \textit{Punishment}, Ch. 5.
intolerably high. Certain preconditions that fall outside of the scope of criminal justice institutions, therefore, might be necessary for the sake of implementing a restorative justice, and hence morally legitimate, approach to the problem of criminal wrongdoing. Like the problems of historic injustice, racism, etc., this is a reminder of how the problem of punishment involves the community, not just as a wronged party where there is crime, but also contributing to the conditions that contribute to the commission of crimes.

As described above, however, a restorative justice theory of punishment grounded in the duties of offenders does allow room for pursuing specific deterrence policies in the special case of psychopathic offenders. Therefore, we should also consider how criminal justice institutions aimed at the ends of restorative justice are equipped to address sociopaths in the general public, i.e. people who have the potential to be psychopaths but have not yet committed any crimes. In my view, we should not view the general public as made up entirely of potential psychopathic offenders. Were it the case that the general public were made up entirely (or even mostly) of sociopaths, i.e. people unable to feel the force of moral reasons not to offend, then we would have good reasons to pursue an institution of criminal punishment aimed at general deterrence rather than the

\[261\] Ibid.

\[262\] For my purposes, sociopaths refer to those who are unable, due to their lack of the requisite psychological capacity, to feel empathy or the force of moral reasons not to offend. By contrast, psychopaths are a subset of sociopaths who are both (1) unable to feel empathy or the force of moral reasons not to offend, and (2) are actually guilty of offending. Many sociopaths are not criminals, although their reasons for refraining from wronging others are not (by definition) moral reasons, but rather merely prudential disincentives of various sorts, e.g. social stigma, the negative effects associated with wronging others to the pursuit of their personal goals, etc.
ends of restorative justice. However, evidence suggests that the vast majority of people are in fact capable of feeling the force of moral reasons. According to Watson, “Approximately 1 percent of the general population worldwide are thought to be [potentially] psychopathic…” Therefore, since the vast majority of people are not sociopaths, a system of coercive criminal law ought to address the general public with the respect they are owed as rational and responsible moral agents capable of responding to moral reasons.

But, one might ask, even if the percentage of the general public that is sociopathic is relatively low, shouldn’t a system of coercive criminal law provide protection against such individuals? How, one might ask, is this to be accomplished within the context of a restorative justice approach? In my view, as previously argued, criminal justice institutions ought to aim at fostering moral recognition in wrongdoers, and reparations between wrongdoers and their victims as co-members of the normative community. As argued in Chapter 4, it is essential to the restorative justice account that wrongdoers are held accountable in some way for their wrongful conduct. The state ought to intervene in the lives of wrongdoers in order to persuade them, through the use of force, to live up to what they owe to others.

Therefore, sociopaths in the general public are provided with prudential reasons not to offend, even if the justifying aim of criminal justice institutions is not general deterrence. If a system of criminal law is clear and determinate, then it allows for the general public, including sociopaths, to predict with a degree of

consistency how they are likely to be treated by the coercive apparatus of the state. As such, for those who are unable to feel the force of moral reasons, a system of criminal justice that aims to provide the general public with relevant moral reasons not to offend, i.e., by holding wrongdoers in some way accountable to the community for what they have done, will be interpreted by certain individuals as providing merely prudential reasons not offend.

Consider the following analogy: suppose that you are my boss, and that I promise to pick you up at the airport. However, after making my promise, I win a free ticket to a concert that I very much desire to attend. The reason I should not fail in my duty to pick you up at the airport, despite my desire to attend the concert, is that it would be wrong to fail in my obligation to keep my promise to you. As a general rule, people should keep their promises to others. Now suppose, in an almost identical scenario, a sociopath is confronted with the same choice: either keep his promise to pick his boss up from the airport, or attend the concert. The sociopath might choose to pick his boss up from the airport for entirely prudential reasons, i.e. to avoid the harm of the negative influence on his career and personal goals associated with failing in his obligation to keep his promise to his boss. Of course, this doesn’t dissolve the moral scenario underlying his choice, i.e. as a general rule, people should keep their promises. Rather, the sociopath simply does the right thing for the wrong reasons. Therefore, while general deterrence is not the justifying aim of a system of criminal law that aims at the ends of restorative justice, from the perspective of
sociopaths in the general public, general deterrence might be a welcome side effect of such a system.\footnote{Such a side effect is “welcome” for the same reasons that we ought to pursue specific deterrence in the case of psychopathic offenders. Sociopaths, like everyone else, have primary obligations to refrain from wronging others.}

### 6.5 – Moral recognition and the legitimate authority of the state

A further objection to the restorative justice account is as follows: one might agree that, as Duff suggests, moral recognition is the appropriate response to wrongdoing on the part of the wrongdoer. However, although moral recognition is required from the perspective of the individual wrongdoer, one might object that it is not the appropriate role of the state to attempt to induce recognition and repentance in wrongdoers through the use of force. One might object by pointing out that the state’s role in the context of criminal justice “…is not that of the mentor or priest.”\footnote{See von Hirsch, Censure and Sanctions, p.10.} In other words, to even attempt to foster moral recognition in wrongdoers would be to go beyond the legitimate purview of the state in the context of criminal justice.

On this point, von Hirsch argues,

> A penance seeks to reach deeper than mere penal censure does: in order to elicit the requisite attitudes of repentance the sanctioner needs to inquire into the person’s feelings – or at least, fashion the sanction so that it is designed to reach those feelings. Might this not be overreaching on the State’s part? Granted it need not be manipulative literally: Duff’s penance is designed only to help the person reach his own penitent understanding, not to compel him to express sentiments he does not feel. Might not the person object, nevertheless, that the penance is an inappropriate form of State intrusion, that while he should be treated as capable of moral
judgement, the nature of his actual attitudinal response to the State’s censure is his own business? 266

Along similar lines, Bennett argues,

It is not part of the remit of the state to pursue the full-blown moral reconciliation that comes with repentance ... Of course the world would be a better place if offenders did respond to expressions of condemnation with genuine repentance and reform. But it is not clear to me that aiming to make this happen is the business of the state, let alone the justification of the criminal sanction. 267

Let this worry be termed the intrusiveness objection. These criticisms of the restorative justice account have inspired philosophers such as von Hirsch and Bennett to defend expressive rather than communicative theories of (retributive) punishment. While both theorists argue that the intrusiveness objection provides decisive reasons to reject the restorative justice account, along with the penal aim of fostering moral recognition in wrongdoers, interestingly, both argue that the criminal sanction should nevertheless leave open the “opportunity” for moral reform on the part of offenders. 268

As I have argued, by contrast, the duties of offenders create limited permissions to harm wrongdoers for the sake of convincing them to reform themselves (i.e. to fulfill their obligation of moral recognition to others in the attitudinal sense). Thus, I agree with Duff that an institution of criminal punishment, if and when such an institution ought to be considered permissible, ought to aim at fostering the appropriate attitudinal response in wrongdoers. Our

266 See von Hirsch, Censure and Sanctions, p.74 (von Hirsch’s italics).
267 See Bennett, The Apology Ritual, p.197.
268 Von Hirsch argues, “Censure gives the actor the opportunity for ... responding [in the appropriate way], but it is not a technique for evoking specified sentiments.” von Hirsch, Censure and Sanctions, p.10.
coercive interventions should aim to persuade wrongdoers to reform themselves, and thereby to foster reparations between wrongdoers and their communities.

To adequately address the intrusiveness objection it is necessary to break down the complaint into its constituent parts.

6.5.1 – The feasibility complaint

The first aspect has to do with a worry concerning feasibility. Von Hirsch points out that it would be exceedingly difficult for the state to aim at moral reform on the part of offenders, since it would involve a considerable degree of investigation into the moral constitution of wrongdoers, their feelings, their respective reasons for offending, etc. In my view, this is the weakest construal of the intrusiveness objection. The fact that morally legitimate coercive interference in the context of criminal justice would be difficult is not, I submit, a good argument against its legitimacy and moral credibility. It is true that Duff’s account and my own present highly normatively ambitious approaches to the problem of punishment, but it is not clear to me why we should agree that since a theory of punishment is ambitious, it is therefore substantively incorrect. For example, contemporary philosophical approaches to the problem of global hunger (or the problem of the moral status of non-human animals, etc.) are highly ambitious in their prescriptions regarding what justice requires. However, rarely do philosophers claim that the normative ambition of such philosophical theories and the difficulties surrounding the implementation of their prescriptions provide decisive reasons to reject their proposals of what justice requires.
6.5.2 – The proportionality complaint

The second aspect of the worry has to do with proportionality. According to von Hirsch, aiming at bringing about moral reform in wrongdoers involves abandoning the important principle of proportional punishment. He argues, Duff’s penance rationale creates an undesirable tension with proportionality. The latter calls for A and B to receive comparably severe punishments, if the gravity of their crimes is approximately the same. The aim of eliciting penitence, however, points the other way: B might have a thicker skin than A, and hence might need a tough penance before the message is likely to penetrate.\(^{269}\)

In the context of von Hirsch’s discussion, this complaint is a species of the charge that it is not the appropriate role of the state to attempt to bring about penitent reform in wrongdoers. As I argued above, however, a proportionality requirement ought to be considered built in to any reasonable account of the secondary obligations of offenders. Furthermore, as I argued in Chapter 4, we should evaluate different modes of expressing the judgement that wrongdoers are blameworthy with reference to the likely forward-looking effects of the mode of expression on the wrongdoer. If a particular mode of expression is more likely than some other mode of expression to bring about penitent reflection in the wrongdoer, then we have a defeasible reason to pursue it. Therefore, while, as a baseline, the severity of our responses to wrongdoing ought to be proportionate

\(^{269}\) See von Hirsch, *Censure and Sanctions*, p.76.
to the seriousness of the offence in question, the mode of expression might nevertheless be tailored toward fostering moral recognition in wrongdoers.\textsuperscript{270}

6.5.3 – Moral recognition and the liberal state

This brings us to the third and final dimension of the intrusiveness objection: von Hirsch and Bennett argue that it is not the appropriate role of the liberal state to meddle in the moral constitution of its citizens. This complaint relates to a deeper challenge to the restorative justice account concerning the legitimate scope of the authority of the state.

Von Hirsch argues,

Answering this objection would require a particular, and fairly ambitious, account of State power – one in which the State not only expresses its disapprobation of certain conduct, but tries to bring about certain responsive attitudes in those whom it condemns. Perhaps such an account can be defended – although I would not be easily persuaded.\textsuperscript{271}

\textsuperscript{270} See section 4, above; von Hirsch, however, claims that this strategy merely illustrates rather than alleviates the tension between inducing penitent reflection and proportional punishment: he writes, “One could, of course, propose that proportionality governs the severity of the punishment and penance its mode: so A and B would receive penalties that have to be comparable in severity, but might differ in other respects so as better to elicit penitence. This, however, just illustrates the tension: such restrictions on severity would make it more difficult to suit the penalty to the actor’s expected responsiveness…” See von Hirsch, \textit{Censure and Sanctions}, p.76. In response, however, I submit that this argument collapses into the worry that morally legitimate criminal punishment is difficult to achieve. See my response to the feasibility complaint above (section 6.5.1).

\textsuperscript{271} See von Hirsch, \textit{Censure and Sanctions}, p.74.
As Bennett approvingly argues, “A properly liberal state … will acknowledge limits on the extent to which it ought to concern itself with the attitudes of the offender.”

While von Hirsch and Bennett agree that the criminal sanction should leave open the opportunity for moral reform, they contend that it should not go so far as to attempt to induce moral reform in wrongdoers. However, I submit that this argument relies on a distinction between having an opportunity in a thin, negative sense – i.e. having an opportunity in the sense of not being prevented from X – versus having an opportunity in a thick, positive sense – i.e. having a substantive opportunity, whereby the agent is enabled in her effort to actually achieve or access X. Recall that on my account as well as the account defended by Duff, moral recognition is not the sort of thing that the criminal sanction seeks to extract from wrongdoers by sheer force. All parties therefore agree that the criminal sanction should provide wrongdoers with the opportunity to recognize the moral gravity of their offences. The difference is that, for von Hirsch and Bennett, this requirement is satisfied merely by not interfering with the wrongdoer’s individual efforts to reform. In my view, however, simply not preventing wrongdoers from achieving moral recognition does not go far enough in providing wrongdoers with substantive opportunities to reform.

---

272 See Bennett, *The Apology Ritual*, p.190
273 See above, note 266. Bennett makes a similar point. He argues, “After all, we cannot force repentance.” See Bennett, *The Apology Ritual*, p.191.
274 This highlights that the language of “inducing” penitent reflection in wrongdoers is misleading. All parties to the debate agree that the criminal sanction should provide opportunities for wrongdoers to recognize the wrongness
Von Hirsch and Bennett therefore agree with the restorative justice account insofar as they admit that the liberal state should provide wrongdoers with the opportunity to reform, despite their deployment of a thin sense of having an opportunity, i.e. having an opportunity to X in the sense of not being interfered with or not being institutionally prohibited from achieving X. For Duff’s account and my own, by contrast, the difference lies in the extent to which, by means of the criminal sanction, wrongdoers are provided with a substantive or reasonably accessible opportunity to reform. Therefore, defending punishment for the sake of moral recognition does not necessarily require an argument for a more expansive or ambitious view of the appropriate scope of the authority of the liberal state. Von Hirsch and Bennett agree that the state should provide the opportunity for offenders to reform themselves. The dispute merely concerns whether such opportunities for moral reform are in fact realistically realizable opportunities.

Furthermore, as previously argued, the secondary obligations of offenders, in my view, supply the justification for why it is not overly intrusive for a liberal state that respects individual freedom to take an active role in persuading wrongdoers to reform themselves. As argued in Chapter 5, wrongdoers, by virtue of wronging others, incur secondary moral obligations toward others. Such obligations include the duty to reform oneself in the requisite, attitudinal sense. The secondary obligations of wrongdoers in turn create limited permissions for a
third party, e.g. the state, to interfere in the lives of wrongdoers to attempt to persuade them – i.e., in the sense of providing wrongdoers with accessible opportunities – to reform. Therefore, while it would indeed be overly intrusive for the state to interfere in the moral constitution of citizens not guilty of criminally wronging others, in the special case of wrongdoing, the duties of offenders create special, limited permissions for the state to interfere.

For example, consider the case of sociopaths. Sociopaths who refrain from offending do not need to justify their moral constitution to the state, and so, subjecting sociopathic citizens to intrusive forms of coercive treatment for the sake of meddling in their moral constitution would be, in my view, impermissible. However, in the special case of wrongdoing, wrongdoers (sociopathic or not) have demonstrated an active willingness to culpably wrong others, and hence ought to be held accountable to the community for their wrongful conduct. By virtue of wronging others, they create for themselves further obligations that are to a greater degree stringent and onerous when compared with the primary obligations we all have to refrain from wronging each other. Therefore, the state is not, in my view, over-stepping in its legitimate authority by interfering in the moral lives of criminal wrongdoers, within certain limits, to persuade them to live up to what they owe to others.

6.6 – Conclusion

In this Chapter, I defended against some prominent objections the view that an institution of criminal punishment ought to aim at the ends of restorative
justice. Andrew von Hirsch argues that such an account fails to provide compelling reasons to hold accountable wrongdoers who are already repentant or predictably defiant. Concerning the already repentant, I argued that such an objection fails to account for the duties of offenders in the attitudinal and expressive senses. It is not enough that a wrongdoer is already repentant; her repentance must be appropriately expressed by a meaningful apology, and must be conditioned by a form of burden that addresses the community to whom her relationship has suffered injury. Concerning the predictably defiant, I argued that even in the face of considerable evidence otherwise, we should not view certain wrongdoers as irredeemably evil. In the more specific case of psychopathic offenders, we have reasons, grounded in the duties of psychopathic offenders, to pursue the aim specific deterrence. Despite psychopathic offenders lacking the capacity for a certain form of moral agency, such offenders nevertheless may be permissibly held accountable for their wrongful conduct, and nevertheless have enforceable obligations to refrain from reoffending.

Since a restorative justice approach allows for the possibility of alternatives to criminal punishment, i.e. criminal punishment understood as one institutional possibility among others for holding wrongdoers accountable and pursuing the ends of restorative justice, I responded to the famous objection surrounding reformatory or corrective alternative modes of criminal treatment, namely, that such an approach allows no room for a principle of proportionality. Critics argue that, for the sake of fairness, coercive forms of criminal treatment must fit the crime. I argued that our use of coercive intervention for the sake of
pursuing the ends of restorative justice ought to be reined in, in both the attitudinal and expressive contexts, by a principle of proportionality. In the expressive context, the burdensomeness of appropriate apology ought to be considered proportionate to the seriousness of the wrongdoing in question. In the attitudinal context, indeterminate modes of criminal treatment would not only be counterproductive, such modes of treatment fail to respect wrongdoers as rational and responsible moral agents. While a non-retributivist approach to moral responsibility would sanction deviations from strict proportionality, since a proportionality constraint is built into any reasonable construal of the duties of offenders, strict proportionality ought to be considered the baseline according to which deviations must be justified.

Subsequently, I responded to Tadros’s objection that pursuing moral recognition via criminal justice institutions would put society at unnecessary risk. For Tadros, according to any theory of punishment that focuses on fostering moral recognition in wrongdoers, “insufficient concern is shown for citizens in protecting them from the harms and wrongs that they might be more or less inclined to suffer depending on whether, how, and whom we punish.” However, in my view, what I call the problem of insufficient protection is, at its basis, an empirical concern. Not only does the restorative justice approach provide room for crime prevention via providing wrongdoers as well as the general public with relevant, moral reasons not to offend, but furthermore, a restorative justice approach can be sensitive to the effects of the particular mode

---

275 See Tadros, *The Ends of Harm*, p.89.
of expressing the judgement that criminal wrongdoers are blameworthy and ought to be held accountable. The efficacy of a particular restorative justice policy in preventing crime is among the legitimate reasons to opt for such a policy in the face of possible institutional alternatives.

Finally, I responded to the intrusiveness objection, i.e. the objection that it is not the appropriate role for the state to aim at bringing about moral recognition in wrongdoers. Von Hirsch and Bennett argue that the state overreaches in its legitimate authority when it attempts to meddle in the moral constitution of offenders. In response, I suggested that the intrusiveness objection relies on a distinction between two senses of “having an opportunity”: if we agree, as von Hirsch and Bennett do, that the state should provide the opportunity for wrongdoers to reform themselves, then the dispute concerns merely to what extent we are willing to provide wrongdoers with accessible or realizable opportunities to live up to what they owe to others. Furthermore, if the worry is really about feasibility, then it fails to be convincing, since many other widely accepted theories of what justice requires in other contexts are exceedingly normatively ambitious. If the worry is really about proportionality, then I suggested that permissible interference, grounded in the duties of offenders, can abide by a principle of proportionality. While the severity of our use of coercion ought to be similar across cases, the mode of expressing the judgement that the wrongdoer ought to be held accountable ought to be sensitive to the likely effects of the mode of expression on the wrongdoer. Finally, if the worry is really about the appropriate scope of state power, then, as I argued in Chapter 5, the duties
of offenders create limited permissions to harm wrongdoers for the sake of
convincing them to live up to what they owe to others as wrongdoers. It would, I
suggested, be impermissible for the state to interfere in the moral constitution of
those who have not, through the act of wronging others, rendered themselves
liable to (limited forms of) coercive treatment. In my view it is permissible,
however, for the state to interfere in the lives of wrongdoers for the sake of
convincing them to live up to what they owe to others in the requisite sense.
Chapter 7 – Conclusion

Criminal wrongdoing damages our relationships with each other as co-citizens. Our response to crime as a community should aim to heal the damage done to our relationships resulting from crime. A restorative justice approach should not presuppose that an institution of criminal punishment is irrelevant or counterproductive to the pursuit of restorative justice. Punishment can be instrumental to fostering recognition in wrongdoers, and enabling them to convey a meaningful apology to the community and those whom they have wronged. Nor should a restorative justice approach assume that an institution of criminal punishment is necessary to holding wrongdoers appropriately accountable to others. Blame can be expressed in a variety of ways. The appropriate mode of expressing blame should involve considerations not only concerning the culpability of the wrongdoer, but also our standing as a community in relation to the wrongdoer, the wrongdoer’s own moral disposition in relation to her wrongdoing, and the likely forward-looking restorative effects of our collective expression of condemnation on the wrongdoer, her victim(s), and the community.

As I argued in Chapter 2, consequentialist approaches to the problem of punishment fail to adequately address two problems: such accounts fail to adequately respect the status of wrongdoers as rational and responsible moral agents, and fail to respect the basic rights of wrongdoers. A restorative justice approach provides solutions to both problems. Restorative justice requires that our responses to crime respect the status of wrongdoers as moral agents.
capable of responding to moral reasons. Wrongdoers have obligations that flow from wronging others. Our response to crime as a community, therefore, ought to build upon what is morally required of wrongdoers as moral agents. Restorative justice also requires respect for the basic rights of wrongdoers. The liability of wrongdoers to be harmed for the sake of pursuing the ends of restorative justice is limited. We should constrain our use of force in the context of criminal justice with reference to what wrongdoers owe to others. The use of force for the sake of restorative justice is permissible only to the extent that it is compatible with the obligations of wrongdoers to recognize their wrongdoing as wrongdoing, and to make amends.

As I argued in Chapter 3, retributivism is a highly controversial approach to the problem of punishment. Despite its weaknesses, however, retributivists are correct in arguing that restorative justice requires a condemnatory response on the part of the community to criminal forms of wrongdoing. But we should resist the more specific conclusion that criminal punishment is always the response that justice requires. Hard treatment is not valuable in itself. Rather, hard treatment is valuable insofar as it contributes to fostering moral recognition in wrongdoers.

Undoubtedly, there are further possible objections to the restorative justice account that, for the sake of brevity, I have not been able to address in this project. Nevertheless, I hope to have provided the groundwork for pursuing further a non-retributive, restorative theory of criminal justice, grounded in the secondary obligations of wrongdoers. If the duties of wrongdoers are taken to
supply the justification for interfering in their lives, then we should take seriously
the implication that an institution of criminal punishment is not the only possible
way for wrongdoers to live up to what they owe to others. However, I recognize
that this view requires further elaboration and clarification, and therefore is likely
to remain, at least for the time being, controversial.
Bibliography


