Thesis title: Social Deprivation and Criminal Punishment

Candidate: Peter Siu Chun Chau, St. Antony's College

Email: pscchau@hku.hk

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in Politics in the Department of Politics and International Relations at the University of Oxford

Date: 13 January 2015

Word count: 98000 words

Thesis Title: Social Deprivation and Criminal Punishment

Abstract:

My aim in this thesis is to examine whether there are some mitigating factors, i.e. reasons to punish an offender less for his crime than an otherwise similar offender (other than that the offender suffered from mental disorder or disturbance or other forms of irrationality at the time of offence)[1], that are more applicable to socially deprived offenders than to non-socially deprived offenders. I will answer the thesis question through a critical examination of twelve arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders, each proposing a different mitigating factor.[2]

My conclusions are as follows:

(1) Most of the arguments that I examine fail, i.e. they either fail to highlight a genuine mitigating factor, or we do not have much evidence that the mitigating factor

---

1 I understand "irrationality" in the sense suggested by Morse. See Morse, "Deprivation and Desert". As an example, the automatism defence in English law is based on the idea of irrationality.

2 Actually, two out of the twelve arguments focus on the same mitigating factor (a bad moral education) but they make very different use of the concept. I will clarify this in chapter 1.

highlighted by the argument has a greater applicability to socially deprived offenders than to non-socially deprived offenders.

(2) However, one argument, which can be called the *no violation of natural duties argument*, is successful.

(3) Moreover, *the improvement of the worst off argument*, an argument that is not often discussed in the literature, is particularly noteworthy. If my discussion about that argument is correct, then even if, as I will argue, the mitigating factor highlighted by that argument may not be more applicable to socially deprived offenders than to non-socially deprived offenders, the remaining parts of that argument would still have profound influence on punishment in our unjust societies.

To Kim-Man, Suet-Ping, Nancy and in memory of Chun-Yung Chan

Acknowledgements

My first thanks go to my supervisors, Dr. Daniel McDermott and Prof. John Gardner. Dr. McDermott has been supervising my work since I began MPhil studies at Oxford (on a different topic). I am very grateful to him for his care and attention in reading my drafts and for his sharp and rigorous criticisms, which have saved me from numerous errors. Prof. Gardner has been an extremely patient and encouraging supervisor. His wide knowledge base and ability to connect ideas have been a constant source of stimulation. He continued to supervise my work even after I relocated from Oxford to Hong Kong; his support exceeded what I could ever ask for.

A deeply felt thank you also goes to a number of other teachers and friends, without the support of whom this dissertation would not have possibly come to fruition. Prof. John Tasioulas was one of my MPhil co-supervisors. He was a model supervisor and was the one who sparked my interest in the philosophy of punishment and walked me through my first steps in this area. Mr. Frank Choi has been reading my work ever since he and I arrived in Oxford. He has been a constant source of encouragement, friendship, and help. He painstakingly reviewed the whole dissertation and offered instructive comments. Prof. Cora Chan also read through several chapters of the thesis and provided most helpful suggestions. Prof. Joseph Chan and Prof. Jeff McMahan have been very supportive since my undergraduate days; without their guidance and encouragement, I probably would not have pursued postgraduate studies at all. Finally, without my family I would not have been able to achieve anything. It follows that without them I would not have been able to finish this dissertation.

My work on this dissertation is supported financially by the Swire Educational Trust. I wish to thank the donors for their generosity.

I have published some materials based on this dissertation. A portion of Chapter 2 is published as "Temptations, Social Deprivation, and Punishment", *Oxford Journal of Legal Studies* (2010) 30(4): 775. Chapter 7 provided the basis for "Duff on the Legitimacy of Punishment of Socially Deprived Offenders", *Criminal Law and Philosophy* (2012) 6(2): 247. Chapter 8 provided the basis for "Poverty, Distributive Justice, and Punishment", *Canadian Journal of Law and Jurisprudence* (2012) 25(1): 39.

Table of Contents

Chapter 1: Introductory chapter

1. Introduction

In this thesis, I aim to examine whether there are some mitigating factors, i.e. reasons to punish an offender less for his crime than an otherwise similar offender (other than that the offender suffered from mental disorder or disturbance or other forms of irrationality at the time of offence), that are more applicable to socially deprived offenders than to non-socially deprived offenders. More specifically, I aim to examine whether there are some descriptions (excluding descriptions that refer to mental disorder or disturbance or other forms of irrationality of the offender) that pass one of the following two sets of tests:

Set 1 (which can apply to descriptions that do not admit of degrees)

a) *The empirical test:* The proportion of socially deprived offenders who satisfy that description is greater than that for those who are non-socially deprived offenders; and

b) *The moral test:* If an offender satisfies that description, then there is a *pro tanto* reason to punish (by "punishment" I mean "criminal punishment" in the rest of this thesis) him less for his crime than an otherwise similar offender (i.e. someone who committed the same crime, caused the same harm with the same degree of foresight, was of the same age, etc.) who does not satisfy that description.

Set 2 (which can apply to descriptions that admit of degrees)

a) *The empirical test*: On average socially deprived offenders satisfy that description *to a higher degree* than non-socially deprived offenders; and

b) *The moral test:* If an offender satisfies that description to a higher degree, then there is a *pro tanto* reason to punish him less for his crime than an otherwise similar offender who satisfies that description to a lower degree.

Different arguments are distinguished by their proposed descriptions. (There is one exception: the description "the offender had a bad moral education" can be a basis for two entirely different *forms* of argument and I will consider them separately. For this point, see section 5 of this chapter below where I discuss chapters 4 and 9.) An argument for claiming that there is a mitigating factor that is more applicable to socially deprived offenders consists of two claims, one empirical and one moral, which correspond to the empirical test and the moral test respectively.

For example, according to the non-culpable normative mistake argument, the proposed description is “the offender acted under a non-culpable normative mistake”; that is, according to this argument, there is a mitigating factor that is more applicable to socially deprived offenders because first, morally, if an offender acted under a non-culpable normative mistake then there is a reason to punish him less for his crime than an otherwise similar offender who did not act under a non-culpable normative mistake, and second, empirically, the proportion of socially deprived offenders who acted under a non-culpable normative mistake is greater than that for those who are non-socially deprived offenders. I regard an argument to be successful if both its empirical claim and moral claim are true.

I will answer the thesis question through a critical examination of different arguments (twelve in total) for claiming that there is a mitigating factor that is more applicable to socially deprived offenders. In doing so, I believe I have covered most prominent arguments (that are not based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence) discussed in the literature, though I make no particular claim to comprehensiveness: I am open to the possibility that there are some sound arguments that are not considered in this thesis.

My answer to the thesis question is that while many arguments discussed in this thesis are unsuccessful, one argument, namely, the *no violation of natural duties argument*, is successful. Moreover, I think the *improvement of the worst off argument,* an argument that is not often discussed in the literature, is particularly noteworthy. I will argue that the moral claim of this argument, contrary to the belief of many, is true. While I cannot claim the argument to be successful because we do not have clear evidence for its empirical claim, if my discussion concerning that argument is correct then even if the empirical claim of that argument is false (in the sense that the mitigating factor highlighted by the argument may not be *more applicable* to socially derived offenders than to non-socially deprived offenders), the remaining parts of that argument would still have profound influence on punishment in our unjust societies.

The structure of the remaining parts of this introductory chapter is as follows. I aim to introduce and motivate my project in sections 2 to 5 of this chapter. In section 2, I will explain my project by clarifying its main terms. In section 3, I will try to justify why I have chosen to undertake this project with its particular scope. In section 4, I will lay out the structure of my whole thesis. In section 5, I will provide a synopsis of the subsequent chapters.

In sections 6 to 9, I will situate my project within the broader literature. In section 6, I will examine the relationship between the arguments considered in this thesis and the recognized defences in English criminal law like self-defence, necessity, duress, and insanity. In sections 7 and 8, I will discuss the relationship between my project and general issues about punishment. I will not attempt to construct a comprehensive theory of punishment in my thesis. But given that this thesis is one about punishment, it is not surprising that I have committed myself to some general claims about punishment in my thesis. In section 7, I will lay out my general view on punishment, which consists of five claims, and state whether, in relation to each of the claims that I have committed to, whether I will defend it or just assume it in my thesis.

Hart has argued persuasively that not all views about punishment are views about *justifications* of punishment, i.e. views about what provides the positive reasons to punish.[3] Sometimes we can commit to a claim about punishment without committing to any claim about what does and does not justify punishment. For example, Hart argued that a person can accept mistake and insanity to be valid criminal defences no matter whether that person believes punishment is justified on retributivist or deterrent grounds.[4] In section 7, I will discuss the issue of what general claims on

3 Hart, "Prolegomenon to the Principles of Punishment".
4 Hart, "Legal Responsibility and Excuses".

punishment I have committed myself to in this thesis; but I will *not*, in that section, explore the issue of what claims on *justifications* of punishment I have committed myself to in this thesis. I will explore the latter issue briefly in section 8.

The concept of mercy has recently attracted more attention in the literature on punishment. In section 9, I will explore which arguments discussed in this thesis are arguments based on mercy. I believe there is no single answer to this question as there are three different conceptions of mercy. I will consider whether the arguments discussed in this thesis are arguments based on mercy on each of the three conceptions of mercy.

2. Clarifying my project

In this section I wish to clarify my project. I want to stress that in this section my aim is to *clarify* my project but *not* to *justify* my project. I will perform the latter task in section 3.

I take social deprivation to mean the lack of social goods. Social goods include not merely material wealth; but also include at least things like education, status, and political power. While there may be disputes about the peripheral cases, we can pick out the central cases of socially deprived offenders in a contemporary Western society, the offenders on whom this thesis will focus, without much difficulty: an offender who grew up in slums, brought up by criminal parents, dropped out of school in third grade, who cannot find any stable employment at the time of offence, etc. certainly can be classified as socially deprived.

I take criminal punishment for a crime to be: (a) a kind of hard treatment that (b) communicates/expresses blame and (c) is imposed by the state.[5] This definition explains, for example, why imprisonment for rape is criminal punishment but tax, social opprobrium, and penance imposed by a person's religious community are not. This definition can, of course, be further clarified and refined. But I think this definition captures most of the normatively significant features of criminal punishment for my purpose and that further refinement of the definition of criminal punishment, while a desirable project in itself, would not affect my discussion.

Concerning the moral test, a "*pro tanto* reason" can be overridden by other considerations.[6] To say that there is a *pro tanto* reason for doing X does not entail that

5 Feinberg, "The Expressive Function of Punishment".
6 On the distinction between *pro tanto* reasons and all things considered judgment, see, for example,

we should do X all things considered, or that X is a very strong reason.

By "The proportion of socially deprived offenders who satisfy that description is greater than that for those who are non-socially deprived offenders" (i.e. the empirical test under set 1), I mean the percentage of socially deprived offenders who satisfy that description is greater than the percentage of non-socially deprived offenders who satisfy that description. For example, if 40% of socially deprived offenders satisfy a description while only 10% of non-socially deprived offenders satisfy it, then I regard it to be true that a greater proportion of socially deprived offenders than non-socially deprived offenders satisfy that description and accordingly that description passes the empirical test. *It is not necessary that all, or most, socially deprived offenders satisfy that description.* Similarly, by "on average socially deprived offenders satisfy a certain description to a higher degree" in the empirical test under set 2, I do not mean that all socially deprived offenders satisfy that description to a higher degree than any non-socially deprived offenders. For example, I regard it possible for the claim "on average socially deprived offenders faced stronger temptations to offend when they offended" to be true even if there are some non-socially deprived offenders who faced stronger temptations than those faced by most socially deprived offenders.

Temkin, "Equality, Priority, and the Levelling Down Objection".

It follows that since, no matter whether we are talking about set 1 or set 2 of the tests, an argument can pass the empirical test *even if not all* socially deprived offenders satisfy the description put forward by the argument (or satisfy it to a higher degree), I do *not* intend to focus *only* on mitigating arguments that can cover *all* socially deprived offenders. So if by saying that "social deprivation is a *distinct* mitigating factor" one means that there is a reason to punish *all* socially deprived offenders less for their crimes than all non-socially deprived offenders, then I will *not only consider* arguments for claiming that social deprivation is a distinct mitigating factor.

3. Justifying my project

3.1. Four questions about my project that need to be answered

In this section I will try to explain why I have chosen to undertake this project with its particular scope by answering the following four questions:

a) The broad choice of the topic*:* Why do I choose to write a thesis on the punishment of socially deprived offenders at all?

b) Moral test: In my thesis, I judge whether an argument is successful depending on whether it puts forward a description that passes both the empirical test and the moral test as *I defined above.* However, one might wonder what the rationale of insisting on these two tests is. Concerning the moral test, why do I formulate it in terms of *pro tanto* reasons to mitigate, rather than *pro tanto* reasons for full defence or all things considered mitigation? In other words, why do I focus on the question of whether there are *pro tanto reasons* to punish socially deprived offenders less for their crimes, but not the question of whether there are *pro tanto* reasons for giving them a full defence or the question of whether all things considered we should punish them less for their crimes?

c) Empirical test*:* Concerning the empirical test, why do I consider an argument passes the empirical test *as long as the proportion* of socially deprived offenders who satisfy the description put forward by the argument is greater than that for those who are non-socially deprived offenders or that on average socially deprived offenders satisfy that description to a higher degree?

d) Exclusion of arguments that are based on mental disorder or disturbance or

other forms of irrationality: In my thesis I investigate whether there are any descriptions that pass one of the two sets of test, *excluding descriptions that refer to mental disorder or disturbance or other forms of irrationality of the offender*. In other words, I will *not* consider arguments that are based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence. What is the rationale for this?[7]

3.2. My answers to the four questions

a) The motivation underlying the broad choice of topic

Many people, at least as a pre-theoretical matter, would think there is something special about socially deprived offenders which often justifies leniency, as compared to offenders who have every advantage in the world but still decided to go against the law. But can these feelings survive rational scrutiny, or will they vanish, after rational examination, as mere sentimental nonsense? When pressed to explain their sympathy

[7] For arguments that appeal to the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence in justifying a reduction of their punishment, see, e.g. Delgado, "'Rotten Social Background': Should the Law Recognize a Defence of Severe Environmental Deprivation?", 38-45; Bazelon, *Questioning Authority*; Lippke, "Diminished Opportunities, Diminished Capacities", 468-471; Wolf, *Freedom within Reason*, 75-76; Morse, "Deprivation and Desert", 143-145; Roberts, "The Ethics of Punishing Indigent Parents"; Falk, "Novel Theories of Criminal Defense Based upon the Toxicity of the Social Environment".

towards socially deprived offenders, common responses include: ‘given their circumstances, what could you expect?’; ‘if we were brought up that way, we would be the same.’; and ‘they have suffered too much already; give them a break.’ Whilst these common responses may embody important insights, they are nonetheless too crude to be doing much justificatory work. A more in-depth treatment of the issue is helpful. Therefore, I regard it to be a philosophically interesting project to consider mitigation of punishment of socially deprived offenders. (This topic, needless to say, is also of some practical importance, given that a substantial proportion of crimes are in fact committed by socially deprived people.[8]) In the last 40 years or so, there have been a number of articles, book chapters, and *obiter dicta* in books that dealt with the topic.[9] However, I am not aware of any book-length monograph specifically on the topic. This thesis is intended to encourage further work in this area, by subjecting various arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders to philosophical scrutiny.

b) Moral test: *Pro tanto* reason for mitigation, not all things considered judgment or full defences

[8] For a review of some relevant empirical data, see Karmen, “Poverty, Crime, and Criminal Justice”, 27-28.

[9] I will not list all of them here, as many of them will be discussed in the coming chapters. George Wright has published a brief book but it has not considered most of the literature being discussed here. See Wright, *Does the Law Morally Bind the Poor?* There is a helpful collected volume of essays on this topic though, which is Heffernan and Kleinig (eds), *From Social Justice to Criminal Justice.*

The rationale of focusing on the narrower question of whether there are some *pro tanto* mitigating factors that are more applicable to socially deprived offenders, rather than the broader question of whether, *all things considered*, those offenders should be punished less for their crimes, is that even the narrower question is theoretically interesting. Only so much can be done in a single thesis. Therefore, I want to proceed step by step by focusing on the narrower issue about *pro tanto* reasons first.

That said, I want to make it clear that while I will focus *primarily* on the *pro tanto* issue, I *will not completely* shy away from the all things considered issue. For example, in chapter 8, after defending the idea that there is a *pro tanto* reason to punish an offender less for his crime if by doing so we can improve the well-being of the worst off members of society, I will explore the question of whether in many cases we should do so, all things considered, by examining whether doing so will violate values that are generally more important than improvement of the worst off.

Similarly, given that even the issue of whether there are some mitigating factors that are more applicable to socially deprived offenders is not uncontroversial, I want to focus on mitigation of punishment rather than full defences first, since if there is not even a case for mitigation there is unlikely to be a case for full defences. However, whilst I will not focus on full defences in my thesis, it is clear that any exploration on the topic of mitigation will have an impact on the topic of full defences. This is because, as Husak plausibly claims, many mitigating pleas are "analogues" of full defences,[10] so there is bound to be some continuity between discussions concerning mitigating pleas and discussions concerning full defences. For example, suppose we accept that offenders who faced stronger temptations to offend when they offended are less blameworthy for their crimes and hence deserve mitigation of punishment.[11] If that is the case, then it is plausible to accept that if the temptations faced by an offender are sufficiently strong, then he is not blameworthy enough to deserve any punishment at all and hence he should be fully excused.

c) Empirical test: Greater applicability

I am looking for arguments that can account for my intuition (which I believe to be shared by many) that there is *something special* about social deprivation in deciding the punishment for an offender's crime. As such, in order for an argument to count as successful, it is insufficient for it to simply highlight a mitigating factor that applies to

[10] Husak, "Partial Defences".
[11] When I use the word "blameworthiness" in my thesis, I mean "moral blameworthiness" unless the context clearly suggests otherwise.

some socially deprived offenders because it seems to me that an argument that shows that social deprivation has some special significance in criminal punishment must *at least* highlights a plea that has *a greater applicability* to socially deprived offenders than to non-socially deprived offenders.

To illustrate this point, consider the following. Some socially deprived offenders would suffer special hardship in prisons. But even if we assume that special hardship in prisons provides a morally valid reason to punish an offender less for his crime, it would be very strange, before further empirical evidence is adduced, to claim that the idea of special hardship shows that there is something special about social deprivation in deciding the punishment for an offender's crime because there is no reason to think that the plea based on special hardship has a *greater applicability* to socially deprived offenders than to non-socially deprived offenders. That explains why I require that, in order for an argument to pass the empirical test, a greater proportion of socially deprived offenders must satisfy the description put forward by the argument than non-socially deprived offenders, or that on average socially deprived offenders satisfy that description to a higher degree.

Another question is why I do not insist on a *more demanding* empirical test. For example, why do I not draft the empirical test in a way such that in order for an argument to pass the empirical test, *all (or at least most)* socially deprived offenders must satisfy the description put forward by the argument or that *all* socially deprived offenders must satisfy that description to a higher degree than non-socially deprived offenders? In other words, why do I not focus *only* on arguments that try to show that there is a mitigating factor that applies to *all* socially deprived offenders?

The reason for not adopting that very demanding empirical test is that I believe it would already be an important discovery if we can find some mitigating factors that have a greater applicability to socially deprived offenders than to non-socially deprived offenders *even if those mitigating factors cannot apply to all socially deprived offenders*. Our intuition about socially deprived offenders seems to be that they, *as a group*, are special. We do not necessarily feel that *all* of them should receive special treatment. Moreover, many arguments in the literature concerning mitigation of socially deprived offenders are not meant to apply to all or most socially deprived offenders.[12] It seems that our project will be over-restrictive if we neglect those arguments by insisting on examining only arguments that are meant to apply to all or most socially deprived offenders.

[12] See, e.g. Buss, “Justified Wrongdoing”; Aquinas, *Summa Theologica*, part 2-2 question 66.

One may object that even if the empirical test should not be tightened to one such that an argument can pass the test *only if all or most* socially deprived offenders satisfy the description put forward by the argument (or satisfy it to a higher degree), the empirical test as I formulated is nonetheless too loose. According to this objection, in order for an argument to justify the idea that there is something morally special about social deprivation in punishment and accordingly to capture our pre-theoretical intuition, it is *not sufficien*t that a greater proportion of socially deprived offenders, *as a statistical matter,* satisfy the description put forward by the argument (or that they on average satisfy it to a higher degree); *in addition, the difference in proportion must be non-accidental*. So for example, according to this objection, even if it happens that in our society 3% of socially deprived offenders acted under self-defence and only 2% of non-socially deprived offenders acted under self-defence, the idea of self-defence does not really justify the idea that there is something morally special about social deprivation in punishment because *the correlation is entirely accidental*. In reply, first, while noting that this objection is quite intuitive, I am not sure whether we should add the non-accidental requirement to the empirical test and what exactly it is. (Is it one about quantity? What if we change the percentage from 3% to 70%?) Second, more importantly, I do not think adding the non-accidental requirement, or tightening up the empirical test in similar ways, would affect my discussion in this thesis substantially. I do not think my acceptance or rejection of the empirical claim of any of the arguments I considered in my thesis depends on whether we add the non-accidental requirement to the empirical test. So if the readers prefer I am happy to add the non-accidental requirement to the empirical test in my thesis and the readers can read in relevant phrases when I discuss the empirical claim of different arguments in the rest of this thesis.

d) Exclusion of arguments that are based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence

There are two reasons for not considering arguments based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence in this thesis. The first reason is that I believe many people feel that there is often a reason to punish a socially deprived offender less for his crime even if he did not act under any kind of mental disorder or disturbance or irrationality. The second reason is that I believe the question of whether and what kind of mental disorder or disturbance or irrationality should ground a reason to punish an offender less for his crime is one of the most difficult and controversial questions in moral philosophy. In order for the other arguments to receive proper attention, I believe it is best to bracket arguments based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence in my thesis.

4. Structure of the subsequent chapters: the four-fold classification of mitigating arguments

Full defences can be divided into three main kinds: justifications, excuses, and non-exculpatory defences.[13]

Both justifications and excuses are *exculpatory* defences, i.e. they are defences that are based on the idea that the offender is not *morally blameworthy* for his crime (or not blameworthy to the extent such that a conviction is warranted). By a person's degree of moral blameworthiness for a certain conduct I, following many writers on this topic, refer to the degree of negative moral *reactive attitudes* he deserves for that

[13] Robinson, "Criminal Law Defences: a Systematic Analysis". While I by and large follow Robinson's structure, I use "justifications" more broadly than him to include both "offense modification defenses" and justifications under Robinson's terminology.

conduct.[14] To say a person's conduct is justified means that his conduct is *not wrongful*, and hence, he is not blameworthy for the conduct. Consent by the victim, self-defence, and necessity are examples of justification. To say a person's act is excused means that he is not blameworthy for the act *despite the fact that his act is wrongful*.[15] Duress is an example of excuse. If a person has a non-exculpatory

---

[14] The idea that moral blameworthiness is connected to deservingness of negative moral reactive attitudes is accepted by many. Many writers on moral responsibility begin their work by accepting that connection and then proceed to consider what makes reactive attitudes deserved. For a few examples, see Strawson, "Freedom and Resentment"; Fischer and Ravizza, *Responsibility and Control*; Wolf, *Freedom within Reason*; Wallace, *Responsibility and the Moral Sentiments*; Rosen, "Skepticism about Moral Responsibility"; Westen, "An Attitudinal Theory of Excuse".

Some people believe that blame should not be analysed in terms of holding reactive attitudes but instead offered two other proposals. First, some people believe that holding some reactive attitudes is not necessary for blame: holding a certain *judgment* is sufficient. Second, some people believe that holding some reactive attitudes is not sufficient for blame: if I simply hold some reactive attitudes against a person without imposing a sanction on him, then I have not blamed him (at least not in the morally important sense). Accordingly, under the second proposal, whether a person is blameworthy depends not simply on whether he deserves reactive attitudes but also on whether he deserves sanctions. (For a discussion of these two proposals of blame and some others, see Smith, "Moral Blame and Moral Protest".) I think these two proposals are mistaken but let me observe that even if one of them is correct, my project would not be severely affected.

If blame is analysed in terms of judgment rather than reactive attitudes, then I believe my position is strengthened rather than weakened. It is because my preferred account on what matters to moral blameworthiness, the "quality of will" account, which I will defend in chapter 4, seems even more plausible if we understand blame as judgment and accordingly blameworthiness as deservingness of moral judgments: it seems very plausible that, if blame is a kind of judgment, then it is a kind of judgment about lack of concern, and accordingly, a person who has less concern for others in committing an act should deserve more blame for that act. (Hieronymi, "The Force and Fairness of Blame". Similar points apply to the proposal, accepted by Smith and Talbert, which understands blame as a form of moral protest. See Smith, "Moral Blame and Moral Protest"; Talbert, "Moral Competence, Moral Blame, and Protest".)

Suppose that blame, contrary to what I think, necessarily involves some kind of sanction. If that is the case, all I need to do to accommodate that is to do some relabeling in my thesis. Let me explain. I divided my thesis into four parts. Parts 1 and 2 are about mitigation of sanction based on the idea that the offender deserves less severe moral reactive attitudes for his crime. Parts 3 and 4 are about mitigation of sanction *not* based on the idea that the offender deserves less severe moral reactive attitudes for his crime. Given my acceptance of the connection between reactive attitudes and blame, I labelled the arguments in parts 1 and 2 "exculpatory arguments", i.e. arguments based on reduced blameworthiness and arguments in parts 3 and 4 "non-exculpatory arguments", i.e. arguments not based on reduced blameworthiness. If, contrary to what I think, blame should be analysed in terms of sanctions, then some arguments in part 3 and 4 should be classified as exculpatory arguments as well (since they bear on the deservingness of sanctions even though they do not bear on the deservingness of reactive attitudes). But the soundness of those arguments and my discussion of them would not be affected by this relabeling – what matters in deciding the soundness of those arguments is still whether they highlight something about the offender that justifies the claim that he deserves a less severe sanction despite the fact that he does not deserve less severe moral reactive attitudes, and whether we call those arguments exculpatory or non-exculpatory would not bear on that question.

[15] Here I use the word "excuse" broadly. Gardner further divides this broad category of excuses into

defence, then it means that he should not be punished for the act *despite the fact that he is blameworthy for the act.* Diplomatic immunity and abuse of process by the prosecution are examples of non-exculpatory defences.

As mentioned, I agree with Husak that mitigating factors are analogues of full defences.[16] Similar to the three-fold classification scheme concerning full defences, we can divide mitigating factors, and accordingly arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders, into three types. First, arguments that claim that there is a partial justification that is more applicable to socially deprived offenders, i.e. arguments that claim that if an offender satisfies the description put forward by the argument then his crime is less wrongful and hence there is a reason to punish him less for the crime. Second, arguments that claim that there is a partial excuse that is more applicable to socially deprived offenders, i.e. arguments that claim that if an offender satisfies the description put forward by the argument then even if his crime is not less wrongful he is less blameworthy for his crime and hence there is a reason to punish him less for the crime. Third, arguments that claim that there is a non-exculpatory mitigating factor that is more applicable to socially deprived offenders, i.e. arguments that claim that if an

exemptions and excuses proper. See Gardner, "In Defence of Defences".

[16] Husak, "Partial Defences".

offender satisfies the description put forward by the argument, then there is a reason to punish him less for his crime even if he is not less blameworthy for his crime.

We can further divide non-exculpatory mitigating factors into two kinds. Punishment is not only hard treatment; it involves communication of blame. So there are two possible bases why an offender should be punished less for his crime despite the fact that he is equally blameworthy for it.

First, it may be that despite the fact that he is equally blameworthy for his crime, there is a reason to punish him less for it because *the right to communicate blame* to him, held by our courts and penal institutions, is compromised. (For simplicity, when I simply refer to "courts" below I should be understood as referring to "courts and penal institutions".) Let us call this type of mitigating factors *non-exculpatory communicative mitigating factors*.[17]

The second kind of non-exculpatory mitigating factors, in contrast, does not focus on blame and the communication of it. If an offender possesses this kind of mitigating factor, then there is a reason to punish him less for his crime, *even if* he is equally

[17] For the most famous proponent of this type of argument, see Duff, *Punishment, Communication, and Community,* chapter 5.

blameworthy for his crime and the right to communicate blame to him held by our courts is not compromised. Let us call this type of mitigating factors *non-exculpatory non-communicative mitigating factors*.

Before I go on, let me observe that while a utilitarian argument for mitigation of punishment (i.e. we punish an offender less because that will promote general utility) is a non-exculpatory non-communicative mitigating argument, non-exculpatory non-communicative mitigating arguments need not take a utilitarian form. First, a non-exculpatory non-communicative mitigating argument need not be consequentialist. For example, Murphy believes that, regardless of further consequences, socially deprived offenders should be punished less for their crimes because of fairness, even if they are equally blameworthy for their crimes and the right to communicate blame to them held by our courts is not compromised.[18] Second, consequentialist non-exculpatory non-communicative mitigating arguments need not be utilitarian.[19] The improvement of the worst off argument and the whole life moral desert argument (which I will discuss in chapter 8 and chapter 10 and which I will introduce in section 5 of this chapter below) are consequentialist arguments not based on blame; but the good consequences that the two arguments appeal to are not simply

[18] Murphy, "Marxism and Retribution".
[19] Sen and Williams, "Introduction".

increase in aggregate utility.[20]

I will divide the main body of this thesis into four parts following this four-fold classification of mitigating arguments. Part 1 of my thesis, consisting of chapters 2 to 5, is concerned with arguments for claiming that there is a partial excuse that is more applicable to socially deprived offenders. In part 2 of my thesis, consisting of chapter 6, I will consider arguments for claiming that there is a partial justification that is more applicable to socially deprived offenders. In part 3 of my thesis, consisting of chapter 7, I will consider an argument for claiming that there is a non-exculpatory communicative mitigating factor that is more applicable to socially deprived offenders. In part 4 of my thesis, consisting of chapters 8 to 10, I will consider arguments for claiming that there is a non-exculpatory non-communicative mitigating factor that is more applicable to socially deprived offenders.

5. Synopsis of subsequent chapters

Part 1: Partial excuses

[20] Scanlon's discussion of punishment also takes into account considerations that affect neither blame nor utility. See Scanlon, "The Significance of Choice".

Chapter 2: Temptation and motivating strength

In chapter 2 I will examine two arguments for claiming that there is a partial excuse that is more applicable to socially deprived offenders. According to the first argument, the *temptations argument*, on average socially deprived offenders *faced stronger temptations to offend when they offended* and an offender is less blameworthy for his crime than an otherwise similar offender if he faced stronger temptations to offend when he offended.[21] In reply to this argument, I argue that the moral claim of this argument is false: it is not true that if an offender faced stronger temptations to offend when he offended then he is less blameworthy for his crime than an otherwise similar offender who faced weaker temptations to offend when he offended.

According to the second argument, the *motivating strength argument*, on average socially deprived offenders *acted on stronger reasons in committing their crimes*, and an offender is less blameworthy for his crime than an otherwise similar offender if he acted on stronger reasons in committing his crime. I will argue that we do not have empirical evidence to accept, even provisionally, the empirical claim of this argument.

[21] See, e.g., Von Hirsch and Ashworth, *Proportionate Sentencing*, chapter 5.

Chapter 3: Prior fault

In chapter 3, I will examine and reject the *prior fault argument* which, unlike the temptations argument and the motivating strength argument, is a historical argument (i.e. an argument that focuses *not on the features of the offender at the time of offence* but on his past). According to the prior fault argument, if an offender is not blameworthy for acquiring the mental states that motivated his crime (i.e. those mental states *are not traceable to his prior faults*),[22] then he is less blameworthy for that crime and accordingly there is a reason to punish him less for that crime than an otherwise similar offender who is blameworthy for acquiring the mental states that motivated his crime (let us call the moral claim of this argument *the prior fault principle*); and the proportion of socially deprived offenders who are not blameworthy for acquiring the mental states that motivated their crimes is greater than that for those who are non-socially deprived offenders.[23] I will argue against the moral claim of this argument.

Chapter 4: Moral education and direct blameworthiness

[22] For the idea of "tracing", see Fischer and Ravizza, *Responsibility and Control*, 49.
[23] See, e.g. Lippke, "Diminished Opportunities, Diminished Capacities", 471.

In chapter 4, I will consider and reject the *education argument*, another historical exculpatory argument. According to this argument, the proportion of socially deprived offenders who had a bad moral education is greater than that for those who are non-socially deprived offenders, and if an offender had a bad moral education, then he is less blameworthy for his crime and accordingly there is a reason to punish him less for the crime than an otherwise similar offender who had a good moral education, *even if the crime by the offender with a good moral education is not (more) traceable to his prior faults.*

In response, I will argue that the moral claim of the argument is false by defending the quality of will view and rejecting the variable expectation view of moral blameworthiness. According to the variable expectation view of moral blameworthiness, a person's moral blameworthiness for a wrongdoing does not simply depend on his quality of will in acting; in addition, it depends on his difficulty (i.e. his capacity and opportunity) to act morally – we lower our expectation of good will for someone who had a greater difficulty to act morally.[24] If the variable expectation view is true, and if offenders with a bad moral education had less opportunity to internalise moral norms and thus to act morally, then the moral claim

[24] E.g. Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances".

of the education argument will be true. However, I will argue that despite its intuitive appeal, the variable expectation view of moral blameworthiness is unmotivated.

Chapter 5: Non-culpable normative mistake

In chapter 5, I will discuss and reject the *non-culpable normative mistake argument*. According to this argument, if an offender acted under a non-culpable normative mistake (i.e. his normative mistake was not culpably acquired) in committing his crime, then he is less blameworthy for his crime (than an otherwise similar offender who did not act under a non-culpable normative mistake) and accordingly there is a reason to punish him less for that crime; and the proportion of socially deprived offenders who acted under a non-culpable normative mistake is greater than that for those who are non-socially deprived offenders.[25] In reply, I will argue that the moral claim of the argument is false.

Part II: Partial justifications

Chapter 6: Natural duties and duty to obey the law

[25] E.g. Buss, "Justified Wrongdoing".

In this chapter, I will consider two arguments for claiming that there is a partial justification that is more applicable to socially deprived offenders. I will defend the first argument and reject the second argument.

The first argument, the *no violation of natural duties argument*, claims that if an offender violated no natural duties in committing his crime, then his crime is less wrongful and hence there is a reason to punish him less for his crime than an otherwise similar offender whose crime involved violation of natural duties; and the proportion of socially deprived offenders who violated no natural duties in committing their crimes is greater than that for those who are non-socially deprived offenders. I believe, as many others do, that this argument is successful. I think the moral claim of the argument is very plausible so I will focus on defending the empirical claim of the argument. I will argue, following Green and Reimen, that while crimes by non-socially deprived offenders generally involve violation of natural duties, a significant group of crimes by socially deprived offenders does not. That group of crime is purely property crimes against rich victims.[26]

[26] Green, "Hard Times, Hard Time", 66-69; Reimen, "The Moral Ambivalence of Crime in an Unjust Society".

The no violation of natural duties argument, though passing the empirical test, can only cover *a particular type* of crimes and hence has a quite limited scope. The second argument which I will discuss in chapter 6, the *general duty to obey the law argument*, is meant to have a more general application. According to this argument, if an offender had a reduced general duty to obey the law then his crime is less wrongful and accordingly there is a reason to punish him less for that crime; and the proportion of socially deprived offenders who had a reduced general duty to obey the law is greater than that for those who are non-socially deprived offenders. I will challenge the empirical claim of the general duty to obey the law argument; in addition, I will argue that even if this argument is successful, it can at most justify a very modest amount of reduction of punishment.

Part III: Non-exculpatory communicative arguments

Chapter 7: Victims of the polity's injustice

In chapter 7, I will consider and reject the *polity's injustice argument*. According to this argument, socially deprived offenders are likely to be victims of the polity's injustice but that is not the case for non-socially deprived offenders; and if an offender is a victim of the polity's injustice, then our courts' standing to communicate blame to him is compromised in a way that their standing to communicate blame to other offenders is not, and accordingly, there is a reason to punish him less for his crime even if he is equally blameworthy for his crime.[27]

In reply, I will argue that the moral claim of the argument is false because it is not true that our courts' standing to communicate blame to an offender is compromised if the offender is a victim of the polity's injustice.

Part IV: Non-exculpatory non-communicative arguments

Chapter 8: Improvement of the worst off

In chapter 8, I will consider the *improvement of the worst off argument*. According to this argument, if by reducing the punishment of an offender we will produce the consequence of improving the well-being of the worst off members of society, then there is an extra reason, based on *distributive justice*, to reduce the punishment of him

[27] Duff, "Principle and Contradiction in the Criminal Law"; Duff, "Whose Law is it Anyway? Inclusion, Exclusion and the Criminal Law"; Duff, *Punishment, Communication, and Community*, 179-201; Duff, *Answering for Crime*, 191-194; Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial". This argument is also hinted to by Bazelon, but Duff is the first one, to my knowledge, who provides an extended defence of the argument. See Bazelon, *Questioning Authority.*

as compared to another offender the reduction of punishment of whom will not lead to that consequence; and reducing the punishment of a socially deprived offender is more likely to produce that consequence than reducing the punishment of a non-socially deprived offender.

Many people reject the moral claim of the argument. They think distributive justice and punishment belong to different spheres so that even if by mitigating the punishment of an offender we can improve distributive justice, there is no reason for a punitive institution to do so. I will argue against that line of thought. In fact, I will argue that not only do we have a *pro tanto* reason to reduce the punishment of an offender if by doing so we will improve the well-being of the worst off members of society; in many cases, we should do so, all things considered, because the values that would be compromised by mitigation of punishment do not enjoy a general priority over the value of improving the well-being of the worst off.

Some may find the empirical claim of the argument quite intuitive. They may believe that reducing the punishment of a socially deprived offender will improve the overall well-being of the worst off members of society (since it benefits the socially deprived offender who is a worst off member of society) while that is not the case for reducing the punishment of a non-socially deprived offender. Yet, I will argue that is too simplistic. While mitigating the punishment of some (not all) socially deprived offenders will produce the consequence of improving the well-being of the worst off members of society, mitigating the punishment of some non-socially deprived offenders will produce that consequence as well. Without clear empirical evidence for the empirical claim of the improvement of the worst off argument yet, it seems premature to accept it.

Even though we may not be entitled to accept the empirical claim of this argument, I believe what is true about this argument will still have profound influence on punishment in our unjust societies. First, as mentioned, the moral claim of this argument is rejected by many. In none of the major Western jurisdictions do judges mitigate the punishment of offenders on the ground that doing so will benefit the worst off members of society. I believe that the improvement of the worst off argument is right in challenging that. Second, although I am not sure whether the mitigating factor, of improving the well-being of the worst off members of society, is *more* applicable to socially deprived offenders than to non-socially deprived offenders, I believe this mitigating factor is applicable to many socially deprived offenders.

Chapter 9: Moral education and burdens

In chapter 4, I will examine whether a bad moral education can form the basis of an *exculpatory* argument, i.e. whether there is a reason to punish an offender who had a bad moral education less for his crime *on the ground that* he is *less blameworthy* for his crime. In chapter 9, I will consider bad moral education *again*, but this time I will consider whether it could form the basis of a *non-exculpatory* argument. According to the *education and burdens argument*, if an offender had a bad moral education then there is a reason to impose less burden on him in punishment than an otherwise similar offender who had a good moral education, *even if they are equally blameworthy for their crimes*; and the proportion of socially deprived offenders who had a bad moral education is greater than that for those who are non-socially deprived offenders.[28] I will reject the moral claim of the argument.

Chapter 10: Whole life moral desert and payment in advance

In chapter 10, I will consider and reject two more non-exculpatory non-communicative arguments. According to the whole life moral desert argument, if

[28] E.g. Litton, "The 'Abuse Excuse' in Capital Sentencing Trials".

by reducing the punishment of an offender we will produce the consequence of promoting whole life moral desert (i.e. the state of affairs where the total level of misery in each person's whole life corresponds to the total level of moral blameworthiness displayed in his whole life), then there is a reason to reduce his punishment that does not apply to another offender the reduction of punishment of whom will not lead to that consequence;[29] and reducing the punishment of a socially deprived offender is more likely to produce that consequence than reducing the punishment of a non-socially deprived offender. I will argue that the empirical claim of the whole life moral desert argument faces similar difficulties as those faced by the empirical claim of the improvement of the worst off argument. Moreover, I will argue that the moral claim of the whole life moral desert argument is less plausible than the moral claim of the improvement of the worst off argument.

According to the payment in advance argument, if an offender's mental states that motivated his crime were formed as a result of a painful process, then such suffering can be regarded as payment in advance for the crime and hence there is a reason to punish him less for the crime; and on average the mental states that motivated the crimes of socially deprived offenders were formed as a result of a more painful

[29] Ezorsky, "The Ethics of Punishment", xxii-xxvii. See also Card, "On Mercy", 182; Goldman, "Toward a New Theory of Punishment", 61.

process.[30] I will reject the moral claim of the argument.

6. The relationship between the arguments discussed in this thesis and defences currently recognized in English law

English law currently recognizes a few defences in the conviction stage. It recognizes consent, necessity and self-defence as justifications; intoxication, factual mistakes, automatism, insanity, duress, diminished responsibility, and provocation as excuses (the latter two are defences to murder only); diplomatic immunity and abuse of process as non-exculpatory defences.[31]

Two questions can be asked about the relationship between these recognized defences and my project. First, assuming that these defences are morally valid, can these defences *directly* justify the claim that there are some mitigating factors that are more applicable to socially deprived offenders? That is to say, is it plausible to believe that the proportion of socially deprived offenders who can successfully raise these defences is higher than the proportion of non-socially deprived offenders who can do

[30] Klein, *Determinism, Blameworthiness, and Deprivation,* 82.
[31] For an overview of these defences, see Ormeod, *Smith and Hogan's Criminal Law: 13th ed*; Choo, *Abuse of Process and Judicial Stays of Criminal Proceedings: 2nd ed*; Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial".

so? Second, even if the answer to the first question is no, do some of the arguments discussed in this thesis share a similar rationale with some of those recognized defences?

While I do not have empirical evidence to answer the first question, which is ultimately an empirical question, I think the answer is probably no. (I will ignore automatism, insanity, provocation, and diminished responsibility since I will not consider, as I have explained above, mitigating arguments based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence.) The scope of the necessity defence is very restricted in English law. It has been held that hunger is not a defence to theft and homelessness is not a defence to trespass.[32] It seems that any but a most trivial proportion of socially deprived offenders can rely on necessity. Once we remind ourselves that it is necessary for an offender to face a threat of death or serious bodily harm, not any lesser form of pressure or temptations, in order to plead duress under English law,[33] it seems that very few socially deprived offenders can plead duress under English law as well. There seems to be no reason to believe that the proportion of socially deprived offenders who can plead consent, self-defence,

[32] *Southwark LBC v Williams* [1971] Ch 734. See also *Re A (Conjoined Twins)* [2001] 2 WLR 480.
[33] *Dao* [2012] EWCA Crim 1717.

intoxication, factual mistake, diplomatic immunity, or abuse of process is higher than the proportion of non-socially deprived offenders who can plead those defences.

But the mere fact that very few socially deprived offenders can *directly* make use of the recognized defences in English law of course does not mean that these defences are entirely unrelated to my discussion in the subsequent chapters. I now wish to explore the question of whether some of these defences share a similar rationale with some of the arguments that I discuss in the subsequent chapters.

Let us start with the justifications recognized in English law. Self-defence is a defence that focuses on the use of force in response to aggression and none of the arguments discussed in my thesis is based on that rationale. The two proposed partial justifications discussed in this thesis are not based on the idea that the offender's conduct prevented some evil or the idea that the victim consented to the conduct, and accordingly, none of the arguments discussed in this thesis shares a similar rationale with necessity and consent.

Let us now look at the recognized excuses in English law, starting with intoxication. Some clarification on the doctrine of intoxication in English law is necessary before we proceed. I agree with Simester that to call intoxication a “defence” or an “excuse” is a misnomer, as the intoxication doctrine in English law is *not* an *exculpation* doctrine but rather an *inculpation* doctrine.[34] This is because it is a device for convicting an offender despite the fact that he did not have the relevant *mens rea* at the time of offence (which is normally required for conviction); intoxication is generally irrelevant in English law if it does not prevent the defendant from forming mens rea.[35] But although there is no intoxication *defence* as such, we can still sensibly refer to the *doctrine* of intoxication and ask if it shares a similar rationale with some of the arguments that I discuss in my thesis.

I believe that the intoxication doctrine shares a similar rationale with the prior fault argument, which I will discuss and reject in chapter 3. The intoxication doctrine is usually said to be based on the prior fault principle.[36] As mentioned above, the prior fault argument is also based on the prior fault principle. So if my criticism of the prior fault principle in chapter 3 is sound, then that would constitute an attack against the intoxication doctrine in English law as well. Discarding the intoxication doctrine in English law is actually not as radical as it sounds. Australia, for example, does not

---

[34] Simester, “Intoxication is Never a Defence”. The English doctrine on intoxication is the “Majewski rule”: see *Majewski* [1976] UKHL 2.
[35] *Kingston* [1994] 3 WLR 519.
[36] Ashworth and Horder, *Principles of Criminal Law: 7th ed*, 194.

have the English intoxication doctrine and there has been consistent call for reform of this area of law in England.[37]

Let us move on to factual mistake. The excuse of factual mistake under English law is that a person will not be convicted of an offence if he made a factual mistake that negates a necessary mens rea element of the offence. For example, intention to kill or to cause grievous bodily harm to a person is a necessary mens rea element of the offence of murder. So if the defendant acted under a factual mistake that the object that he was shooting was inanimate, he would not have the necessary intent and thus will not be convicted of murder. The non-culpable normative mistake argument, which I will discuss and reject in chapter 5, is also about mistakes. Does the excuse of factual mistake in English law share a similar rationale with the non-culpable normative mistake argument?

Despite their apparent similarities, they do not share a similar rationale for two reasons. First, the factual mistake defence in English law is available even if the mistake is acquired culpably (barring acute intoxication cases covered by the intoxication rule).[38] In contrast, the non-culpable normative mistake argument tries to

---

[37] *O'Connor* (1980) 29 ALR 44; Ashworth and Horder, *Principles of Criminal Law: 7th ed*, 202-204.
[38] See, e.g., *R v C* [2013] EWCA Crim 223.

justify mitigation only for offenders who acquired the mistaken belief non-culpably. Second, under English law only factual mistakes (which negate mens rea) can excuse. Normative mistake, even if non-culpably held, is generally not a defence under English law. In fact, I will argue in chapter 5, based on the quality of will account of moral blameworthiness, that the differential treatment offered to factual mistakes and normative mistakes under English law is morally appropriate.

Let us move on to duress. I believe that one of the two arguments discussed in chapter 2, the motivating strength argument, shares a similar rationale with the defence of duress. The rationale of the duress defence is that when there is a sufficiently strong reason from self-interest for an agent to commit crime and he succumbed to that reason, then he is not sufficiently blameworthy to be sanctioned by the criminal law. The motivating strength argument is based on a similar rationale as it also looks at the strength of reasons that the offender acted on: it claims that there is a partial excuse that is more applicable to socially deprived offenders because they are more likely to act on stronger reasons than non-socially deprived offenders in committing their crimes.

Since I will not consider arguments that are based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence in this thesis, the arguments discussed in the subsequent chapters do not share a similar rationale with automatism, insanity, diminished responsibility, or provocation. (That said, I will discuss some issues related to exemptions and provocation in chapter 4 when I lay out my preferred account of moral blameworthiness.)

Now let us move on to non-exculpatory defences recognized in English law, starting with abuse of process by the prosecution. One difficulty in deciding whether this defence shares a similar rationale with any of the arguments discussed in my thesis is that it is unclear what exactly is the rationale of the abuse of process defence.

Some writers suggest that the rationale of the abuse of process defence is that because of the misconduct of the polity, the courts' standing to communicate blame to the offender is compromised.[39] If that is the case, then the abuse of process defence shares a similar rationale with the polity's injustice argument, which I will discuss in chapter 7. I will attack the polity's injustice argument by arguing that the courts' standing to communicate blame is not affected by the misconduct of the polity. If my

[39] Duff, "A Criminal Law for Citizens"; "Blame, Moral Standing, and the Legitimacy of the Criminal Trial".

discussion concerning the polity’s injustice argument is correct, then I will have shown that one of the possible justifications for the abuse of process defence does not work. Of course, that on its own will not show that the defence is not justified. Perhaps the defence can be justified on other grounds, like the good consequences (e.g. deterrence of prosecutorial misconduct) produced by recognizing the defence.

Other non-exculpatory defences recognized in English law, like diplomatic immunity, do not seem to share much with the non-exculpatory arguments discussed in this thesis. None of the those defences is based on the idea of improvement of the position of worst off, promoting whole life moral desert, payment in advance, or the idea that there is an objection against imposing a burden on someone who had a bad moral education.

7. Commitment to general claims on punishment

As mentioned in section 1 above, I will *not* articulate a comprehensive theory of punishment in my thesis. But since this thesis is one about punishment, I of course have committed to some general claims about punishment in it. Some of my commitments should be obvious after going through the thesis structure and synopsis in sections 4 and 5 above; regardless it may be worthwhile to list the five general claims on punishment that I committed to in my thesis here. (My general position, which consists of these five claims, can be called "*upper-bound pluralistic retributivism"*.)

(1) Proportionality (under both cardinal and ordinal interpretations) is a valid sentencing principle, i.e. there is a *pro tanto* reason to punish proportionately.[40] By proportionality, I mean that the severity of punishment imposed for a crime should be proportionate to the offender's *blameworthiness* for the crime.

(2) In particular, we should make the severity of punishment imposed for a crime proportionate to the offender's *direct* blameworthiness for the crime instead of his *indirect* blameworthiness for the crime.

(3) While cardinal proportionality sets both upper and lower limits to punishment based on the offender's blameworthiness for the crime, the upper limit is a very strong one and much stronger than the lower limit.

(4) The four traditional sentencing considerations, i.e. proportionality, deterrence, rehabilitation, and incapacitation, are relevant considerations in punishment; but they do *not* exhaust all relevant considerations in punishment.

[40] For the distinction, see von Hirsch, "Proportionality in the Philosophy of Punishment".

(5) Punishment is *not* justified as a measure to communicate blame on behalf *of the whole polity*.

I will now go through the five claims one by one and explain, in relation to each of them, where I have committed myself to it and whether I will *defend* it or just *assume* it.

7.1. Claim (1): Proportionality as a valid sentencing principle

I accept the proportionality principle, i.e. there is a *pro tanto* reason to make punishment for a crime fits the offender's moral blameworthiness for the crime. There are two places in my thesis where I commit myself to this principle.

First, in part 1 (chapters 2-5) and part 2 (chapter 6) of my thesis, I will explore exculpatory arguments, i.e. arguments that claim that there is an exculpatory mitigating factor that is more applicable to socially deprived offenders. In examining the exculpatory arguments, I will assume proportionality and I will focus only on examining whether, for the argument that I am considering, that argument highlighted a factor that can reduce the offender's blameworthiness for his crime and that has a greater presence in socially deprived offenders. For example, in examining the argument based on temptations, I will only focus on examining whether an offender who faced stronger temptations to offend when he offended is less blameworthy for his crime, and whether on average socially deprived offenders faced stronger temptations to offend when they offended, without questioning the principle of proportionality.

Second, in chapter 8 I will claim that crime rate is not always sensitive to the severity of sentence in a reasonably just sentencing system (i.e. a sentencing system that is constituted by sentences that are appropriate *before* we consider mitigation based on improving the well-being of the worst off). One of my arguments for that claim is that since proportionality is a valid sentencing principle, a reasonably just sentencing system will be at least partly structured by proportionality; and crime rate will not always be sensitive to the severity of sentence in a sentencing system that is at least partly structured by proportionality.

I will not defend proportionality and I will *just assume* it to be a valid *pro tanto* sentencing principle in my thesis. I do not want to pretend that my assumption is uncontroversial: by assuming the truth of cardinal proportionality, which implies that there is a *lower* limit of punishment set by the offender's blameworthiness for his conduct, I have assumed the truth of positive retributivism, the idea that there can be a positive reason to punish an offender even if no further good consequences can be produced by doing so.[41] The idea of positive retributivism of course has been rejected by some writers.[42] Yet, it should also be observed that many contemporary writers on punishment accept positive retributivism.[43] Since any project needs to start somewhere I will take the truth of the proportionality principle as the starting point.

Instead of defending this assumption directly, I will instead note two points to make it less controversial. First, while I assumed proportionality, I am open to how exactly it can be justified. Some people believe that proportionality can be defended by appealing to the idea of restoring a fair distribution of benefits and burdens prior to the commission of crimes;[44] some people believe that it can be defended by appealing to expressive or communicative considerations;[45] while others believe that it can

---

[41] For the distinction between "positive" and "negative" retributivism, see Duff, *Punishment, Communication, and Community*, 12.
[42] For critique of positive retributivism, see, e.g. Dolinko, "Some Thoughts about Retributivism"; for some recent arguments for the claim that criminal punishment should not depend on moral blameworthiness at all, see Kelly, "Culpability and Fault in Criminal Justice"; Shoemaker, "Blame and Punishment".
[43] Alexander and Ferzan, *Crime and Culpability*, 6-9; Morse, "Deprivation and Desert", 119-121; Duff, *Punishment, Communication, and Community*, chapter 4; Moore, *Placing Blame*; Husak, "'Broad' Culpability and the Retributivist Dream"; Berman, "Punishment and Justification"; Brink, "Responsibility, Incompetence, and Psychopathy".
[44] Murphy, "Marxism and Retribution"; Morris, "Persons and Punishment".
[45] Duff, *Punishment, Communication, and Community;* Kleinig, "Punishment and Moral Seriousness".

simply be defended by appealing to our intuitions about desert.[46] I do not believe I need to take any stance in the debate between these views. Second, by assuming that there is a reason to punish proportionately, one does not need to also assume that such a reason is a lexically trumping one. It is entirely consistent with my assumption of proportionality that proportionality should sometimes be sacrificed for other considerations, like utility, in punishment.[47]

7.2. Claim (2): We should make the severity of punishment imposed for a crime proportionate to the offender's *direct* blameworthiness for the crime instead of his *indirect* blameworthiness for the crime

It is common, in the literature on moral responsibility, for writers to draw a distinction between a person's direct blameworthiness for a conduct and his indirect blameworthiness for a conduct.[48] A person's indirect blameworthiness for a conduct depends on to what extent that conduct was traceable to other prior conducts for which he is directly blameworthy and how directly blameworthy he is for those

[46] Moore, *Placing Blame.* For a survey of those theories, see Boonin, *The Problem of Punishment*; Ryberg, *The Ethics of Proportionate Punishment*, chapter 1. For a sophisticated discussion, not about these different theories of retributivism as such, but on the possible underlying structures of them, see Berman, "Two Kinds of Retributivism".

[47] Berman, "Two Kinds of Retributivism"; Husak, "Retributivism in Extremis", 12-13; White, "Pro Tanto Retributivism".

[48] E.g. Smith, "Culpable Ignorance", 563.

conducts. To illustrate the distinction, consider the following two persons. The first person recklessly took a drink at t1, which made him unconscious at t2. He attacked someone under the unconscious state at t2. The second person was identical except that he was forced to take the drink at t1. On the widely agreed assumption that voluntariness when performing a conduct is a necessary condition for a person to be directly blameworthy for that conduct, both of them are not directly blameworthy for attacking at t2. But the first person, unlike the second, is indirectly blameworthy for that.

In the literature, much attention has been devoted to the following two issues: first, whether proportionality is a valid sentencing principle has been hotly debated in the literature on punishment. Second, what affects a person's direct blameworthiness for a conduct has been hotly debated in the philosophical literature on free will and moral responsibility.[49] But less attention has been paid to the issue of whether, assuming that proportionality is a valid sentencing principle, whether the principle should count both direct and indirect blameworthiness or rather should count only direct blameworthiness.

[49] Two debates loom large. The first one is the traditional debate of whether moral blameworthiness is consistent with causal determination. The second one is about whether a person's direct blameworthiness for an act depends on his mental states at the time of acting only or rather would depend on his history. (The second debate is different from the first since there can be historical compatibilists.)

I will take a side in the last issue and commit myself to the claim that an offender's *indirect* blameworthiness for a crime should *not* be taken into account in deciding the punishment for the crime. I will not just assume, but *will positively argue* for this claim in chapter 3 in two steps. First, I will argue that *if* separate convictions and punishments have been imposed on the agent for his prior blameworthy conducts leading to his present crime, *then* there would be *no* reason to take his *indirect* blameworthiness for the present crime into account in deciding the punishment *for the present crime*; otherwise we would be double counting his prior blameworthy conducts. Second, I will argue that, if some prior faults, which caused the present crime, should be taken into account in criminal punishment, then we should, because of expressive considerations, impose separate convictions and punishments *for those prior faults instead of* taking them into account in deciding the punishment *for the present crime*.

I wish to note one point about my stance towards expressive considerations here. As I mentioned in the last paragraph, I will appeal to expressive considerations to explain why, if we want to take some prior faults leading to a present crime into account in punishment, we should impose separate convictions and punishments for those prior faults instead of taking them into account in deciding the punishment for the present crime. But I have also stated, in section 7.1 above, that I will *not* commit myself to the claim that expressive considerations justify proportionate punishment. Is there any inconsistency here?

The answer is no. The issue of whether expressive considerations should affect the *total amount of punishment* an offender receives is a controversial one: many people think that there is no reason to express blame by burdensome means like punishment if there are less burdensome means to express blame.[50] But I am using expressive considerations in a much less controversial manner here: I am arguing that we can take into account expressive considerations in choosing between two systems of punishment that will inflict *the same total amount of punishment but in a different manner,* namely, a system that takes prior faults that caused the present crime into account in deciding the punishment for the present crime and a system that does *not* take those prior faults into account in deciding the punishment for the present crime but which imposes separate convictions and punishments for those prior faults.

[50] Scanlon, "The Significance of Choice", 214.

7.3. Claim (3): Cardinal proportionality sets a strong upper limit on punishment

Besides committing to the principle of cardinal proportionality, the principle that an offender's blameworthiness for his crime sets (*pro tanto*) upper and lower limits to the amount of punishment that should be imposed for the crime, I also commit to some kind of *asymmetry* in the principle: that the upper limit is much more stringent than the lower limit.

I commit to this asymmetry in chapter 8 where I discuss the improvement of the worst off argument. There I argue that if by *reducing* the punishment of an offender we would improve the well-being of the worst off members of society, then in many cases we should do so all things considered. I then considered an objection against my position that claims that I am forced to accept, as a matter of consistency, the counter-intuitive claim that in many cases we should aggravate the punishment of offenders for their crimes, all things considered, if that would result in an improvement of the well-being of the worst off members of society. One of my rejoinders to that objection is that going *above* the appropriate sentence set by the offender's blameworthiness for his crime is much more difficult to justify than going *below* it.

The asymmetry claim is not very controversial. Many writers who accept proportionality accept this asymmetry.[51] The asymmetry claim can be defended in various ways. For example, one may defend it, as Morris does, by appealing to the idea of rights: if we go above the upper limit in punishing an offender we violate his rights but it is hard to see how we can violate his rights by going below the lower limit.[52] I will *assume* this asymmetry in the proportionality principle in my thesis without further argument.

7.4. Claim (4): The four traditional sentencing considerations, i.e. proportionality, deterrence, rehabilitation, and incapacitation, are relevant considerations in punishment; but they do not exhaust all the relevant considerations in punishment

Many academic works on sentencing suggest that in deciding the appropriate gravity of punishment, we have to pay attention to the four traditional sentencing considerations: proportionality, deterrence, rehabilitation, and incapacitation. Sometimes a certain work would suggest that some of the four are irrelevant, or that

[51] E.g. Tonry, "Proportionality, Parsimony, and the Interchangeability of Punishments", 231-232.
[52] Morris, "The Future of Imprisonment", 1173.

one of the four should enjoy a dominant or primary status.[53] But it is not frequently suggested that a consideration other than the four should be taken into account in deciding the quantum of punishment.

I think these four considerations are all relevant considerations in punishment. The idea that the three traditional consequentialist considerations (deterrence, rehabilitation, and incapacitation) can sometimes affect gravity of punishment is not controversial – in any case I will *assume* them to be relevant sentencing considerations. (Yet, let me also observe that none of the main points in my thesis depends on this assumption.) I have also stated, in section 7.1 above, that I will assume that proportionality is a relevant sentencing consideration.

However, I disagree with the claim that these four considerations *exhaust* all the relevant considerations in punishment since I will claim, in chapter 8 of my thesis, that distributive justice, in the sense of improving the well-being of the worst off members of society, a consideration distinct from the four traditional sentencing considerations, should be a relevant consideration in deciding the gravity of punishment. I will *defend this claim positively* in chapter 8.

---

[53] See, e.g., Ashworth, *Sentencing and Criminal Justice: 5th ed*. Recently, the idea of "restoration" is more frequently mentioned. However, I will argue that there are values relevant to punishment besides the four traditional considerations *and* restoration. See also *Criminal Justice Act 2003* s142.

7.5. Claim (5): Punishment is not justified as a measure to communicate blame on behalf of the whole polity

As stated above in section 7.1, while I assume the truth of proportionality I will not commit myself to a particular way of defending proportionality. That said, I will commit myself to *ruling out one particular way* of defending proportionality in chapter 7 of my thesis.

Duff suggests that proportionate punishment can be justified as a measure to communicate the appropriate amount of blame on behalf *of the whole polity*.[54] If Duff's suggestion is true, and if it is also true (as I will agree) that the right to communicate blame to socially deprived offenders held by the polity as a whole has been compromised,[55] then there would be an argument for treating socially deprived offenders differently in punishment. However, in response to Duff's argument, I will claim that punishment is *not* justified as a measure to communicate blame on behalf of the whole polity. In other words, I will claim that the right of our courts to punish a person is not compromised just because the right to communicate blame to him held

[54] Duff, *Punishment, Communication, and Community*.
[55] Duff, *Punishment, Communication, and Community,* 179-201

by the polity as a whole is compromised.

I will *argue for this claim positively* in chapter 7. Here I will just note that my disagreement with Duff is not huge. While I disagree with Duff's *particular communicative justification* of punishment, my arguments there would *not* rule out communicative justifications of punishment *in general*. I actually will work within a communicative framework in that chapter. The point I raise there is that the courts need not be seen as communicating blame *on behalf of the whole polity*, as Duff insists, but instead can be seen as communicating blame on behalf of *a subset of the polity, namely, the just citizens*.

(As a matter of caution, I wish to state that while my discussion in chapter 7 is *consistent with* a communicative justification of punishment, it does not *require* a communicative justification of punishment. It is entirely consistent with my discussion there that proportionality, or punishment in general, is justified entirely on non-communicative grounds.)

7.6. Other commitments in my thesis?

I have just stated the five general claims on punishment that I commit myself to in my thesis. I believe that they, by and large, are the general claims on punishment that I commit myself to in my thesis.

There are of course many other general claims that I commit myself to in my thesis, but they are *not,* at least not directly, *claims about punishment*. For example, in chapter 4, I will commit myself to a claim about (direct) moral blameworthiness by arguing that a person's moral blameworthiness for a conduct (i.e. the degree of negative moral reactive attitudes he deserves for that conduct) depends on his quality of will in performing the conduct. I will commit myself to some views about the requirements of distributive justice in chapter 8. I will also make some empirical claims in my thesis (e.g. crime rate is not always sensitive to the severity of sentence).

Those claims can be true or false and I will try my best to argue for them in my thesis. Whether my effort is successful is a question left for the readers to judge after reading the whole thesis. But it should be observed that the truth or falsity of those claims is largely independent from traditional debates in punishment. For example, the issue of whether or not there is a reason to punish proportionately and whether proportionality trumps other considerations in punishment are to a large extent independent from the issue of what is the proper account of blameworthiness, the issue of what are the requirements of distributive justice, and the issue of whether crime rate is always sensitive to the severity of sentence.

8. Commitment to claims on justifications of punishment

In section 7, I have stated what general claims on punishment I have committed myself to in this thesis. Since not all claims about punishment are claims about *justifications* of punishment (i.e. claims about what does and does not provide the positive reasons to punish),[56] it may be worthwhile to state explicitly what claims about justifications of punishment I have committed myself to in my thesis.

By committing to cardinal proportionality in this thesis, I have committed to the claim that retribution is one of the justifications of punishment. As stated in section 7.1 above I will not defend this claim but will just assume it. It should be noted that, however, by doing so I have not committed myself to two other more controversial claims. First, as stated in section 7.1 above, I need not commit to a particular version of retributivism. Second, I need not commit to the idea that retribution is the *sole*

[56] Hart, "Prolegomenon to the Principles of Punishment".

justification of punishment: it is entirely consistent with my view that there are other justifications of punishment like deterrence.

I have stated in section 7.4 above that I will assume that traditional consequentialist sentencing considerations, namely, deterrence, rehabilitation, and incapacitation, are valid considerations in punishment and I will treat them as such in chapter 8. I will assume that they are valid justifications of punishment as well (though I wish to state that none of my main points in the thesis depends on this assumption).

9. Mercy

Mitigating arguments are sometimes classified into arguments based on mercy and arguments that are not. Which of the arguments discussed in this thesis are mercy-based arguments?

In order to answer this question, we need, of course, a clearer understanding of the concept of mercy. One of the central ideas of mercy is that it is *to be contrasted with justice*. But *not every mitigating argument that is not based on justice is an argument* *based on mercy*. Mercy is grounded on *individualistic concern for the offender*.[57] As such, consequentialist mitigating arguments that appeal to the production of good outcomes that *need not involve benefit to the offender himself* cannot properly be called mercy-based arguments. For example, suppose I mitigate an offender's punishment because another person threatened to kill everyone if I do not do so. In such a case, the mitigation is not based on justice. But it does not seem to be based on mercy either.[58]

Most people who discussed the concept of mercy agree that mitigating arguments based on mercy are to be contrasted with both mitigating arguments based on justice and consequentialist arguments that appeal to the production of good outcomes that need not involve benefit to the offender himself. However, three conceptions of mercy can be found in the literature as we can distinguish between *three different conceptions of justice.*

The first group of writers regard mercy as mitigating considerations not based on *legal justice.* For them, mercy is about mitigation of punishment *that is not based on*

57 Card, "On Mercy", 186-187; c.f. Smart, "Mercy", 353-354.
58 For this example, see Card, "On Mercy", 186-187.

the legal rules.[59] An example of this kind of mercy is when the rules provide for a sentence of three months but the judge decided to give a one month sentence because of the special circumstances of the offender (say, the offender acted under duress and mitigation for duress is *not* provided for in the legal rules). Let us call the conception of mercy held by the first group of writers *equity mercy*.[60]

The second group of writers regard mercy as mitigating considerations not based on *censuring justice,* i.e. considerations not based on the idea that the offender *deserves* less censure (i.e. communication of blame) from the punisher. Let us call the conception of mercy held by the second group of writers *pluralistic mercy*.[61] On this conception of mercy, if an offender is given a lower level of punishment for his crime than what he ought to receive based on how much censure he deserves from the punisher for his crime, then that is mitigation based on mercy no matter whether or not such mitigation is provided for in the legal rules; and if an offender is given the punishment that is appropriate based on how much censure he deserves from the punisher, then that is not an exercise of mercy even if the written rules provide for a

---

[59] A different, but for my purpose a similar, position is that mercy is about giving the offender a less severe punishment than "what would normally be expected in a case of this sort". Dolinko, "Some Naïve Thoughts about Justice and Mercy", 354.

[60] This is what Tasioulas referred to as the "reductive conception" of mercy. See Tasioulas, "Mercy"; Johnson, "Entitled to Clemency".

[61] This is what Tasioulas referred to as the "pluralistic conception" of mercy. See Tasioulas, "Mercy"; Smart, "Mercy"; Card, "On Mercy".

higher level of punishment. For example, suppose our legal rules state that *life imprisonment* is the sentence for petty theft. A judge, ignoring the legal rules, sentenced a petty thief to a small fine and let us suppose that, morally speaking, a small fine actually reflects the proper amount of censure the petty thief deserves from the court. This judge is acting mercifully under the equity conception (since he imposed a sentence lower than that based on the legal rules) but he is not acting mercifully under the pluralistic conception (since he did not impose a sentence lower than that based on how much censure the petty thief deserves from the court).

The third group of writers regard mercy as considerations not based on *any kind of justice*, i.e. not based on the concept of rights of the offender at all. On this conception of mercy, which can be called the *radical* conception of mercy, as long as a mitigating argument claims that the offender *has the right* to be punished less for his crime, that argument is not based on mercy.[62] According to this conception of mercy, the central idea of mercy is that the offender cannot require it to be given and the punisher did not do the offender any wrong in denying it from the offender. To illustrate this conception of mercy, let us, for illustrative purpose only, assume the truth of Murphy's unfair advantage theory of punishment and his idea that, as a matter of

[62] Hestevold, "Justice to Mercy".

justice, poor people who received less benefit from society deserve less punishment for their crimes. If we punish a poor offender less because of his disadvantage, it is an exercise of mercy under the pluralistic conception of mercy since he does not deserve less censure for his act; but it is not an exercise of mercy under the radical conception of mercy since the poor offender has a right to less punishment.[63]

I do not think there is much point in deciding which of the three conceptions of mercy is the correct one since that issue seems to be a verbal dispute to me. What I will do, instead, is to go through the three conceptions of mercy and state which of the arguments discussed in my thesis would be classified as a mercy-based argument under each of the three conceptions of mercy.

None of the arguments that I consider in this thesis are based on equity mercy. Equity mercy only has a place in the *law-application* stage; it is about *how judges should respond to sentencing rules*. All of the arguments considered in my thesis, in contrast, are not about how judges should respond to sentencing rules. They are rather arguments about *what sentencing rules we should have in the first place*.

[63] Murphy, "Marxism and Retribution".

Which arguments in my thesis are arguments based on pluralistic mercy? The arguments discussed in parts 1, 2, and 3, are not arguments based on pluralistic mercy since they all try to establish the conclusion that socially deprived offenders often, *as a matter of right*, *deserve less censure* from the courts. So let us focus on the four arguments discussed in part 4 of this thesis.

The improvement of the worst off argument, while *not* based on censuring justice, is *not* an argument based on pluralistic mercy. It is because the concern of that argument, which is consequentialist, is to produce the good outcome of improving the well-being of the worst off members of society, an outcome that can be achieved *even if the offender himself did not benefit from it*. To illustrate, the idea of improvement of the worst off members of society would support mitigating the punishment of a deprived offender if by doing so we can improve the situation of other deprived people (say, his family) even if the offender himself does not benefit from the mitigation of punishment. The same point applies to the whole life moral desert argument, which is also a consequentialist argument.

In contrast, the education and burdens argument and the payment in advance argument are based on pluralistic mercy: they are both arguments, based on an individualistic concern for an offender, for giving the offender a lower level of punishment for his crime than what he ought to receive based on how much censure for his crime he deserves from the court.

Which arguments discussed in my thesis are arguments based on radical mercy? Arguments that are based on censuring justice are clearly based on justice, so arguments discussed in parts 1 to 3 of my thesis are not based on radical mercy. So again we can safely focus on the arguments in part 4.

The improvement of the worst off argument and the whole life moral desert argument are not arguments based on radical mercy. This is because, as I just mentioned, both arguments are not based on an individualistic concern for the offender.

The education and burdens argument and the payment in advance argument, while not based on the idea of censuring justice, seem to be based on the idea of rights. For example, Klein refers to the idea of justice in putting forward the payment in advance argument.[64] Litton, in arguing that an offender who had a bad moral education ought to receive a lesser burden in punishment, appeals to the idea of fairness.[65] So it seems

[64] Klein, *Determinism, Blameworthiness, and Deprivation*, 169.
[65] Litton, "The 'Abuse Excuse' in Capital Sentencing Trials", 1064.

to me that both arguments are not based on radical mercy.

10. Published papers

I have published three papers based on this thesis. Chapters 2, 7, and 8 draw heavily on those three papers:

(1) Chau, P. (2010) “Temptations, Social Deprivation, and Punishment”, *Oxford Journal of Legal Studies* 30(4): 775

(2) Chau, P. (2012) “Poverty, Distributive Justice, and Punishment”, *Canadian Journal of Law and Jurisprudence* 25(1): 39

(3) Chau, P. (2012) “Duff on the Legitimacy of Punishment of Socially Deprived Offenders”, *Criminal Law and Philosophy* 6(2): 247

Part I: Partial Excuses

Chapter 2: Temptations and motivation

1. Introduction

In this chapter I will assess two arguments, namely, *the temptations argument* and *the motivating strength argument*, for claiming that there is a partial excuse that is more applicable to socially deprived offenders.

In section 2, I will introduce and clarify the temptations argument. An assessment will follow in section 3. I will focus on arguing that the moral claim of the temptations argument is false. In section 4, I will respond to a rejoinder made by Lippke in which he defends the temptations argument. In section 5, I will discuss the motivating strength argument, which emerged from our discussion in section 3. I will argue that, at present, we do not have sufficient evidence to accept the empirical claim of the motivating strength argument.

2. Introducing the temptations argument

Let me lay out the two main claims of the temptations argument, using the framework that I introduced in chapter 1, before discussing it:

<u>The Temptations Argument</u>

Moral Claim: If an offender faced stronger temptations to offend when he offended, then he is less blameworthy for his crime and accordingly there is a reason to punish him less for the crime than an otherwise similar offender who faced weaker temptations to offend when he offended. (I will call this claim, a claim about the mitigating force of temptations, the "temptations account of mitigating factors".)

Empirical Claim: On average, socially deprived offenders faced stronger temptations to offend when they offended.

Let me begin by making three brief clarifications on what von Hirsch and Ashworth, two of the most prominent advocates of the temptations argument,[66] mean by

[66] It should be observed that while von Hirsch and Ashworth state that there is a reason to punish an offender who faced stronger temptations to offend when he offended less for his crime, they deny that temptations affect the offender's "culpability". See von Hirsch and Ashworth, *Proportionate Sentencing*, 68, 73. Accordingly, it may be argued that I misconstrued the temptations argument by treating it as an exculpatory argument. To this challenge I wish to make two replies.

First, most other advocates of the temptations argument seem to take the temptations argument to “temptations”.[67] (I will focus on the temptations argument as formulated by von Hirsch and Ashworth, though it should be observed that a similar argument has been advanced by many other writers.[68])

First, the idea of temptations refers to the *agent’s prudential reasons*. A person faces stronger temptations to do X if he has a stronger prudential reason to do X. For example, a hungry person faces stronger temptations to eat than a non-hungry person, as he has stronger prudential reasons to eat.[69]

---

be an argument about blameworthiness. For example, see Lippke, “Diminished Opportunities, Diminished Capacities”, 480 n38; Hudson, “Beyond Proportionate Punishment”. It is, in fact, hard to understand why the strength of temptations faced by an offender when he committed his crime should matter to his punishment if it does not affect his blameworthiness for the crime. The only reason why von Hirsch and Ashworth classify the temptations argument as an argument based on “compassion” rather than “culpability” relies on an implausible and uncommon understanding of the idea of culpability, under which an offender is less culpable for a crime only if it would be “heroic” for him to refrain from committing the crime. See von Hirsch and Ashworth, *Proportionate Sentencing*, 68.

Second, for my purpose it does not matter whether the temptations argument is understood as an exculpatory argument as opposed to a non-exculpatory mitigating argument. This is because the thought experiment (the overdetermined killer case, which I will discuss in section 3) I relied on in suggesting that an offender’s blameworthiness for his crime does not depend on the strength of temptations he faced *also seems to suggest that the amount of punishment an offender should receive for his crime does not depend on the strength of temptations he faced*. To get a bit ahead of myself, I believe that most people, when they reflect on the thought experiment, would feel *not only* that Daisy is equally blameworthy as Devil; *in addition, they would feel that there is no reason to impose a different punishment on them*.

[67] These clarifications are intended to be brief, since I believe that, up to a point, disputes about the concept of temptations will not substantively affect my discussion below. For two quite different analyses of the idea of temptations, see Day, “Temptation”, 175; Lyons, “Coercion as Temptation”, 35. See also section 4 of this chapter below.

[68] E.g. Lippke, *Rethinking Imprisonment,* 84-88; Frank and Groves, “Punishment, Privilege, and Structured Choice”; Hudson, “Punishing the Poor”; Hudson, “Beyond Proportionate Punishment”, 65-71; “Punishment, Poverty, and Responsibility”, 583; Odudu, “Retributive Justice in an Unjust Society”, 427; Wright, *Does the Law Morally Bind the Poor?*, 66; Mann and Al-Khadha, “Freedom of the Will and Criminal Culpability”. For related points see also Hart, “Prolegomenon to the Principles of Punishment”, 15, 50-51; Kant, *Lectures on Ethics,* 62-63.

[69] Von Hirsch and Ashworth, *Proportionate Sentencing,* 63. Some proponents of the temptations account of mitigating factors may insist that for reasons to count as temptations, the agent must, in addition to having those reasons, desire to comply with such reasons. For the distinction between desiring X and having a reason to do X, see Watson, “Free Agency”. I do not think granting this will affect any of my discussion below.

Second, the strength of temptations to do X that a person faces depends on his *net* prudential reasons for doing X; accordingly, it depends *both* on his prudential reasons for doing X *and his prudential reasons against doing X*. If a person has weaker prudential reasons against doing X, then other things being equal he faces stronger temptations to do X. To illustrate, suppose that Helen and Henry equally want to eat a cake without paying. But there is a police standing next to Henry so if he eats the cake without paying he is likely to get caught, while if Helen does so it is likely that she can get away with it. In such a case, Helen faces stronger temptations to eat the cake without paying.[70]

Third, in saying that a person faced temptations to do X, must we imply that he believed himself had *conclusive* reasons *against* doing X, such that he would be acting *akratically* if he really did X, as suggested by Lippke?[71] I have found nothing in the literature by von Hirsch and Ashworth, or in that of the majority of advocates of the temptations argument, to suggest that actions prompted by temptations imply akrasia. My assessment of the temptations argument will begin without the assumption that actions prompted by temptation imply akrasia. Whether or not the temptations argument would be rendered more plausible by accepting Lippke's

[70] Von Hirsch and Ashworth, *Proportionate Sentencing,* 66-68.
[71] Lippke, "Social Deprivation as Tempting Fate", 287. See also section 4 of this chapter below.

narrower understanding of the concept of temptations will be examined in section 4 of this chapter.

3. Assessing the temptations argument

Von Hirsch and Ashworth gave three justifications for the empirical claim of the temptations argument. First, the goods obtained by committing a crime will often be less valuable to a non-socially deprived person. Take car theft as an illustration. A non-socially deprived person may easily afford a car; not so for the socially deprived person; so the car is more valuable to the socially deprived person.[72] Second, criminal punishment generally affects the career of a non-socially deprived person more than a socially deprived person. A street-cleaner's career may not be greatly affected by a term of imprisonment; not so for a lawyer. Therefore, non-socially deprived people generally have stronger prudential reasons to stay away from crime.[73] Third, criminal conviction generally affects the social life of a non-socially deprived person more than a socially deprived person. Criminal conviction generally carries a great stigma in the middle and upper classes. For socially deprived persons it may not be a big deal to be convicted, since many people around them have had that experience; for some

[72] Von Hirsch and Ashworth, *Proportionate Sentencing*, 67.
[73] Von Hirsch and Ashworth, *Proportionate Sentencing*, 67.

of them, it may even be a symbol of status.[74] So again, since the consequences of conviction are more serious for a non-socially deprived person than a socially deprived person, socially deprived people generally face stronger temptations to offend.[75] I will not challenge the empirical claim of the temptations argument here. (Though I will challenge it when I discuss the next argument, the motivating strength argument, in section 5.)

Here, I want to focus on the moral claim of the temptations argument, i.e. the temptations account of mitigating factors. Von Hirsch and Ashworth have not explained why the temptations account of mitigating factors should be accepted. But I guess the appeal of the temptations account of mitigating factors is not hard to find. I believe that many people who accept the temptations account of mitigating factors do so because they think it is supported by our intuition concerning thought experiments similar to the one provided by the following two cases.

First case, tempted killer (Derek): Derek was told by a gangster that if Derek did not kill a person the gangster will cut off Derek's toe. Derek would not have killed if his toe was not at stake. However, caring (too much) about his toe, he finally gave in

[74] Von Hirsch and Ashworth, *Proportionate Sentencing,* 67.
[75] Von Hirsch and Ashworth, *Proportionate Sentencing,* 67.

and killed.

Second case, untempted killer (Devil): Devil killed someone. Unlike Derek, Devil did not receive any threats and his toe was not at stake. Derek and Devil were identical in all morally relevant aspects unless otherwise specified.[76]

Devil and Derek did not act reasonably in the circumstances, and thus are blameworthy for killing.[77] But I think most people believe Derek is *less* blameworthy for killing than Devil. The consideration of these two cases seems to lend some support to the temptations account of mitigating factors. Derek indeed faced stronger temptations to offend than Devil when he offended. It may thus be natural to think that Derek is less blameworthy for killing *because* he faced stronger temptations to offend when he offended, and to generalise from this case and think that the fact that an offender faced stronger temptations to offend when he offended would imply that he is less blameworthy for his offence, other things being equal.

I believe that the temptations account of mitigating factors is false. I wish to start by drawing a distinction between the *temptations account of mitigating factors* and the

[76] Similar cases can be found in Husak, "Partial Defences", 183-184; Frankfurt, "Coercion and Moral Responsibility", 37.
[77] This is to rule out the possibility that Derek has a justification or an excuse.

motivating strength account of mitigating factors. The motivating strength account of mitigating factors is the following principle: If A acted on stronger temptations than B in offending, then A is less blameworthy than B if they are otherwise similar. The strength of temptations an offender acted on in offending is the strength of temptations without which the offender would not have offended.[78]

The distinction between the temptations account of mitigating factors and the motivating strength account of mitigating factors will be understood once we notice that 'the strength of temptations one faced when one offended' and 'the strength of temptations one acted on in offending' need not be the same, and accordingly, the related point that 'A faced stronger temptations to offend than B when they offended' does not entail that 'A acted on stronger temptations in offending than B'.[79] Let's say a person faced temptations of strength *m* for him to offend when he offended. But it

[78] I have defined "the strength of temptations that one acted on" roughly as "the strength of temptations without which one would not have so acted". To be strictly correct, this definition needs to be suitably amended to accommodate for the "Frankfurt Cases" ("Jones$_4$" in Frankfurt, "Alternative Possibilities and Moral Responsibility", 6-8) and to accommodate for other problems related to deviant (i.e. non-rational) causal chain cases. For one illustration of "deviant causal chain", see Davidson, *Essays on Actions and Events*, 79. The amendment can be carried out in various ways, but I will not undertake the task of spelling out the detailed amendment here. Suffice to say that my present preference is to follow generally the proposal offered in Fischer and Ravizza, *Responsibility and Control*, chapters 2 and 3.

[79] The distinction is quite similar to the distinction between "having a reason to do X when one did X" and "acting on that reason in doing X". The distinction between "the reason one acted on" and "the reason one had" can be found in Davidson, *Essays on Actions and Events,* 9, 232, but the idea of analysing phrases like "the reason one acted on" in counterfactual terms is taken mainly from the work of Duff. See Duff, *Intention, Agency, and Criminal Liability*, 58-60. While I disagree with Brandt's ideas that a motivation needs to be "standing" to be relevant for punishment and that excuses in law and morality can be explained by utilitarianism, I find his "A Utilitarian Theory of Excuses", "A Motivational Theory of Excuses in the Criminal Law", "Traits of Character", and "The Structure of Virtue" (all in his *Morality, Utilitarianism, and Rights*) very influential to my thinking about the issue.

may be possible that he would have offended anyway as long as he faced temptations of strength *n*,[80] where *m>n*. *If this is the case, then n, not m, is the strength of temptations that he acted on in offending.* He did *not* act on the strength of temptations represented by (m-n) *despite* the fact that he faced those temptations when he offended. We can call such strength of temptations the *surplus strength*, since that amount of temptations is unnecessary for the offence – the agent did *not* offend *because of* the surplus strength of temptations.

After drawing the distinction between the motivating strength account of mitigating factors and the temptations account of mitigating factors, let us revisit the thought experiment involving Derek and Devil (i.e. the killer cases). Our intuition concerning that thought experiment is that Derek is less blameworthy for killing than Devil. Many people believe that if we accept that intuition, then we are compelled to accept the temptations account of mitigating factors, since they think that the temptations account of mitigating factors is the *only* explanation for that intuition. But once we grasp the distinction between the temptations account of mitigating factors and the motivating strength account of mitigating factors, it is clear that that intuition does not compel us to accept the temptations account of mitigating factors, as the motivating

[80] I.e. *n* is the necessary strength of temptations required for him to commit that offence at that time.

strength account of mitigating factors can also explain that intuition. Derek did not only *face* stronger temptations to kill than Devil when they killed. Derek also *acted on* stronger temptations in killing. As stated in the scenario above, Derek would not have acted but for the fact that killing could save his toe – Derek acted because he wanted to save his toe. Devil, on the other hand, did not act on the temptations to rescue his toe in acting (since his toe was not at stake at all).

I believe it is not hard to construct a thought experiment that will lend strong support to the idea that we should accept the motivating strength account of mitigating factors and reject the temptations account of mitigating factors. Consider the following *overdetermined killer case*, which is a modified version of a case originally given by Harry Frankfurt:

Third case, overdetermined killer (Daisy): Daisy is the same as Derek in all morally relevant aspects, except that while the gangster threatened Daisy that he would cut off Daisy's toe if Daisy did not kill, *Daisy would have killed anyway even if she did not receive the threat*. Knowing that killing could happen to save her toe, she did what she would have done anyway, i.e. killing the victim.[81]

[81] Frankfurt, "Alternative Possibilities and Moral Responsibility", 4. (The example I modified on is Jones$_3$ in Frankfurt's article.) See also Frankfurt, "Coercion and Moral Responsibility", 41. Examples

I believe most of my readers will share my intuition that Daisy is not less blameworthy for killing than Devil. This piece of intuition is consistent with the motivating strength account of mitigating factors but inconsistent with the temptations account of mitigating factors. Daisy indeed faced stronger temptations to kill than Devil when she killed – so if the temptations account of mitigating factors is true then she should be less blameworthy for killing than Devil. On the other hand, the strength of temptations that Daisy acted on was the same as that Devil acted on. Daisy did not act because of the extra temptations – the extra strength of temptations provided by her toe was only surplus strength. Therefore, the motivating strength account of mitigating factors is consistent with saying that Daisy is equally blameworthy for killing as Devil.

Some people may find piecemeal appeals to intuition in moral debates unsatisfying. As such, they would be reluctant to give up the temptations account of mitigating factors just because of one single thought experiment. I think the point, that we should accept the motivating strength account of mitigating factors but reject the temptations account of mitigating factors, can be defended on grounds independent of the thought

similar to the overdetermined killer case can be found in, e.g., Richards, "Acting under Duress", 32-35.

experiment involving Daisy. The fact that A acted on stronger temptations in violating a moral norm (say 'do not kill') than B is constitutive of the fact that A would only be willing to give up the compliance with the norm for stronger prudential reasons than B, i.e. A had more concern for the norm (or the values underlying the norm) than B.[82] Other things being equal, A's motive in offending is therefore better, or less bad, than that of B. And since it is plausible to believe that a person's motive in doing an act should be relevant to his blameworthiness for the act, it follows that the strength of temptations that an agent acted on should be relevant to his blameworthiness for that act. (I will defend this quality of will view of moral blameworthiness in greater detail in chapter 4 of this thesis below.) On the other hand, it is unclear why the mere fact that an agent faced stronger temptations to offend when he offended should reduce his blameworthiness for the offence. If an offender is more blameworthy for an act, then the act should reflect more badly on him. But the strength of temptations an agent faced when he acted does not reflect on his personality if the agent did not act on that amount of temptations.[83]

---

[82] Here I draw on ideas from Brandt, *Morality, Utilitarianism and Rights,* 242-245; Hoffman, "Aquinas on Threats and Temptations", 228, 238.

[83] The point that it is motive, not the strength of temptations, that matters is noticed by Husak in a brief passage in Husak, "Partial Defences", 189-190. See also Richards, "Acting under Duress", 32-35; Maitland, "The Relation of Punishment to Temptations", 262; Bentham, *An Introduction to the Principles of Morals and Legislation,* 149-150 (chapter XI, section XLII, rule 3). (It should be noted that Bentham's terminology is slightly different.) For writings expressing similar points in the context of justifications for wrongs, see e.g., Simester, "Why Distinguish Intention from Foresight", 91-98; Cryer and Simester, "Iraq and the Use of Force: Do the Side-effects Justify the Means", 38-41.

My diagnosis as to why the temptations account of mitigating factors is so alluring, despite being false, is similar to Frankfurt's diagnosis as to why many people find the 'principle of alternative possibilities' convincing.[84] It is true that in ordinary transactions, we would find a wrongdoer less blameworthy for an act as long as he utters 'I faced strong temptations to do that'. But, to borrow Frankfurt's point in a similar context, the reason for us accepting such utterances as valid mitigating factors is only because we assume that most wrongdoers, in making such an utterance, are not 'being disingenuous'.[85] In other words, we assume that a wrongdoer would not utter, as a plea of mitigation, 'I faced strong temptations to do that', unless he also believed that he would have refrained if he did not face so strong temptations. In cases where the assumption holds, the strength of temptations one faced when he offended affects his motive in offending. The fact that we make that assumption in ordinary cases may account for why many people find the temptations account of mitigating factors initially attractive. But we should not lose sight of the fact that what ultimately does the mitigating work is the strength of temptations that the agent acted on, not the strength of temptations one faced to offend when he offended *per se*.

[84] In addition to borrowing ideas from Frankfurt, I model the structure of this paragraph closely on a paragraph by Frankfurt. Frankfurt, "Alternative Possibilities and Moral Responsibility", 9.
[85] Frankfurt, "Alternative Possibilities and Moral Responsibility", 9.

4. Rejoinder on behalf of the temptations account of mitigating factors by Lippke

Lippke published a paper, entitled "Social Deprivation as a Tempting Fate", which is partly a critique of my objection against the temptations account of mitigating factors. He firstly argued for a different understanding of the idea of temptations. Based on that, he then made two main claims against my critique of the temptations account of mitigating factors. *The negative claim* is that my thought experiment involving Daisy and Devil is not a counterexample to the temptations account of mitigating factors, since on a proper understanding of the idea of temptations it is not true that Daisy faced stronger temptations to kill than Devil. *The positive claim* is that the temptations account of mitigating factors can be defended by appealing to the idea of self-control.

As I have mentioned in section 2 of this chapter above, my analysis above proceeds on what can be called a *broad conception of temptations*. On this conception, whether a person faced temptations to do X and how strong the temptations are depend *simply* on his net prudential reasons for doing X. On the *narrow conception* of temptations proposed by Lippke, in contrast, a person faced temptations to do X *only if* (a) he has some prudential reasons to do X; *and also (b) he believes himself to have, all things considered, conclusive reasons against doing X*. The difference between the two conceptions is that the broad conception allows for the possibility of a person who faced temptations to do X and who then did X non-akratically. Under the narrow conception, however, if a person faced temptations to do X and he did X, then the action must be akratic. As Lippke puts it,

> "[We might draw a distinction between two understandings of temptations.] Call [the first] understanding of temptation Temptation1 (=S is tempted to do X if S desires to do X and has no compelling reasons for not doing X). The problem with Temptation1 is that it does not capture what is most familiar about the phenomenon of temptation. Usually when we say that someone is tempted to do X, we have in mind someone ...who has a strong desire to do X but also has good reasons of some kind against doing it. Think about cases of spouses tempted to have extra-marital affairs or citizens tempted to cheat on their taxes. The desires to do these things might be powerful, even overwhelming. But the individuals involved have reasons—prudential or moral or both—against acting on their desires. Call these paradigmatic cases of temptation Temptation2 (=S is tempted to do X if S desires to do X but has compelling reasons for not doing X)."[86]

[86] Lippke, "Social Deprivation as Tempting Fate", 281.

I do *not* want to settle the question of, conceptually speaking, which conception of temptations is closer to our ordinary understanding of the idea of temptations. I am interested in the normative question of whether there is a partial excuse that is more applicable to socially deprived offenders. Allow me, therefore, *simply to address* the question of whether understanding temptations in the narrow sense, as suggested by Lippke, can rescue the temptations argument.

I think understanding temptations in the narrow sense cannot rescue the temptations account of mitigating factors. Let us examine Lippke's negative claim first. Lippke's argument for his negative claim is that although Daisy had stronger prudential reasons to kill, since she did not regard herself as having conclusive reasons against killing she did not face stronger temptations to kill than Devil. Hence, while Daisy may be equally blameworthy as Devil, it is not a counterexample to the temptations account of mitigating factors. As Lippke puts it,

> "[The] initial misstep in his understanding of temptation haunts Chau throughout his analysis...What seems very odd about [Chau's discussion concerning Daisy and Devil] is the suggestion that Daisy faced more

> temptation to kill Victim than Devil. This seems correct only if we focus on the fact that she was threatened. But the threat seems to have left her totally unfazed. It did not suddenly make killing Victim seem attractive to her though she was strongly inclined against it. Like Devil, she had her own sufficient reasons for killing Victim. Daisy does not appear to have been internally torn about what to do and so does not suffer paradigmatic temptation (that is, Temptation2)...In short, I would say that Daisy faced no stronger temptation to kill than Devil, though she was threatened."[87]

In reply to Lippke, I think Lippke cannot avoid the main point of the thought experiment involving Daisy and Devil, and thus rescue the temptations account of mitigating factors, by utilizing a narrow conception of temptations. The main point that I want to get across using the thought experiment involving Daisy and Devil is that if a portion of temptations *played no role in causing* the agent's act, then it should not be taken into account in considering her blameworthiness for her act. Whether or not Daisy acted akratically is of no relevance in the thought experiment. However, in order to directly respond to Lippke's objection, let us imagine two cases where akrasia is unquestionably involved:

[87] Lippke, "Social Deprivation as Tempting Fate", 281.

Fourth case, tempted akratic killer (Eric): Eric had a strong desire to have a car. By killing his victim, he could acquire a car. He thought that all things considered he should not kill. Yet, he succumbed to the desire and killed his victim.

Fifth case, overdetermined akratic killer (Elaine): Elaine had a strong desire to have a car and a moderate desire for an ice cream. If she killed her victim, she could obtain *both* an ice cream and a car. She thought that all things considered she should not kill. Yet, she succumbed to her desires and killed her victim. Furthermore, we know that Elaine's moral fortitude was so weak that she would have killed even if, by killing, she would have obtained *only* the car.

Most of us would believe that Elaine is not less blameworthy for killing than Eric. But this verdict is inconsistent with the temptations account of mitigating factors on Lippke's understanding of the idea of temptations since Elaine faced stronger desires to kill and she acted akratically in killing.

Lippke's narrow conception of the temptations account of mitigating factors is not only counter-intuitive. It is also theoretically unmotivated. This becomes evident when we examine Lippke's positive claim. The claim is that the temptations account of mitigating factor is theoretically attractive because punishment should take into account how much self-control has the offender exercised.[88]

This purported theoretical justification of the temptations account of mitigating factors faces a serious problem. The problem is that an offender who faced stronger temptations to offend when he offended, under the narrow conception of temptations, is *not* necessarily an offender who exercised a higher level of self-control (which is an analogous point to a point I made in section 3: an offender who faced stronger temptations to offend when he offended is not necessarily an offender who acted on stronger temptations). So even if we assume, for the sake of argument, that self-control is relevant to blameworthiness and punishment such that if an offender exercised a higher level of self-control then he should be blamed and punished less for the crime, it does not follow that the strength of temptations an offender faced when he offended, *per se*, is relevant to blame and punishment. The example of Eric and Elaine above aptly illustrates this point. Eric and Elaine faced different degrees of temptation, but they exercised the same level of self-control.

[88] Lippke, "Social Deprivation as Tempting Fate", 284-286.

For the reasons stated above, I believe we should reject the temptations account of mitigating factors, *even if* we accept the narrow understanding of the concept of temptations proposed by Lippke. Quite independently of this, I wish to add the observation that the temptations argument may have very limited applicability if we accept the narrow understanding of the concept of temptations. Taken in its narrow sense, the temptations argument covers *only* socially deprived offenders who committed their crimes akratically. If we believe that actions are frequently non-akratic, bearing in mind that historically many writers have argued that akrasia is not even possible,[89] then Lippke's argument can justify mitigation for only a very small number of socially deprived offenders.

5. The motivating strength argument

In the last two sections I rejected the temptations argument by rejecting the temptations account of mitigating factors. However, my rejection of the temptations account of mitigating factors in favour of the motivating strength account of mitigating factors may contain the seeds of a new argument for claiming that there is a partial excuse that is more applicable to socially deprived offenders. Before

89 For a survey, see Stroud, "Weakness of Will".

discussing it, I will present the two main claims of this new argument, which can be called *the motivating strength argument.*

The Motivating Strength Argument

Moral Claim: If an offender acted on stronger temptations in offending, then he is less blameworthy for his crime and accordingly there is a reason to punish him less for the crime. (This claim is the motivating strength account of mitigating factors, which I discussed in section 3 of this chapter.)

Empirical Claim: On average, socially deprived offenders acted on stronger temptations in offending than non-socially deprived offenders.

As my discussion in section 3 indicates, I accept the moral claim of the motivating strength argument. But I do not think we are entitled to accept the empirical claim of the motivating strength argument, which is ultimately a claim about empirical psychology, even provisionally, at present.

Given that we do not have a comprehensive survey of the psychology of all offenders, we do not have direct evidence for the empirical claim. Of course, the lack of direct empirical evidence, *per se*, is not fatal to an empirical claim (including the empirical claim of the motivating strength argument). Sometimes we can make a plausible suggestive argument for an empirical claim, i.e. an argument that provides us with reasons to accept the claim before we have direct empirical evidence against the claim. If we can never hold empirical claims, at least provisionally, without direct empirical evidence, making sensible policies would be virtually impossible.

The issue, then, is whether we have sufficient indirect support for the empirical claim of the motivating strength argument. I believe the answer is no. I will consider one suggestive argument for the claim and I will argue that it is not sufficiently supported, given our present state of knowledge. Moreover, I will argue that there is a suggestive argument against the claim, which we do not yet have sufficient empirical evidence to refute.

The suggestive argument for the empirical claim of the motivating strength argument starts with the uncontroversial first claim that social deprivation statistically correlates with crime strongly.[90] The argument then puts forward the second claim that socially deprived offenders, on average, faced stronger temptations to offend when they

[90] See, e.g., Karmen, "Poverty, Crime, and Criminal Justice", 26-28.

offended. For the sake of illustration, let us ignore matters of degree and divide temptations into two types, strong and weak. The second claim can then be restated as that the majority of socially deprived offenders faced strong temptations to offend when they offended while the majority of non-socially deprived offenders faced weak temptations to offend when they offended.

The argument then continues by alleging that if both claims are true, then the majority of socially deprived offenders would not have offended if they faced only weak temptations to offend, i.e. they acted on strong temptations in offending. This is because the majority of socially deprived offenders faced strong temptations to offend when they offended, and *most of those offenders, given the strong statistical correlation between crime and social deprivation, would not have offended if they instead belonged to a different social group (i.e. non-socially deprived persons) the members of which generally face weak temptations to offend*. In contrast, the majority of non-socially deprived offenders must have acted on weak temptations in offending. (Since the strength of temptations one acted on is the minimum level of temptations required for one to offend, one could not have acted on stronger temptations than one actually faced in offending.) Accordingly, there seems to be a plausible case for claiming that on average socially deprived offenders acted on stronger temptations in offending.[91]

This argument is quite intuitive. But as it stands, I have two reservations about this argument. My first objection is that even if both claims put forward by the argument are true, it does not follow that the majority of socially deprived offenders would not have offended if they faced only weak temptations to offend. This is because even if most of the socially deprived offenders who faced strong temptations to offend when they offended would not have offended if they are not socially deprived and social deprivation strongly correlates with the strength of temptations one faced when one offended, it does not follow that most of the socially deprived offenders who faced strong temptations to offend when they offended would not have offended if they faced only weak temptations to offend.

An analogy may clarify this point. Suppose that most people with a high IQ do not offend. Suppose further that low IQ strongly correlates with a skin problem. Even if both claims are true, it does not follow that most offenders who have a low IQ and who have the skin problem would refrain from committing their crimes if they did not have the skin problem since low IQ may correlate with some other property P that

[91] For this argument and my subsequent discussion of it, I am indebted to Lewis, "Temptation, Stratification, and Proportional Punishment".

causes crimes, and property P may not strongly correlate with the skin problem. Analogously, if there are, as suggested by many criminologists, other possible causal mechanisms where social deprivation may increase crimes besides increasing the temptations to commit crimes that people face (for example, social deprivation may increase crimes not only through increasing temptations but through its cultural effect on people's concern for moral norms),[92] then the conclusion of this suggestive argument does not follow readily from the truth of the two claims put forward by the argument.

My second objection against this suggestive argument is that while its second claim may, at the end of the day, be true, I do not think it is backed by sufficient empirical evidence. I have not seen empirical evidence for the claim that socially deprived offenders faced stronger temptations to offend when they offended on average. Von Hirsch and Ashworth, as we have seen in section 3 of this chapter above, provided some support for the idea that socially deprived *people* faced stronger temptations to offend on average. That can be granted but it does not follow that on average socially deprived *offenders* faced stronger temptations to offend when they offended since we

[92] This hypothesis, which appeals to culture and respect for moral norms, is supported by the fact that poor Asians are less involved in violent crimes than poor persons of other ethnicity. For this point and a survey of a number of possible ways in which social deprivation may cause crimes, see Karmen, "Poverty, Crime, and Criminal Justice", 27-30. See also Markowitz, "Socioeconomic Disadvantage and Violence"; Markowitz and Felson, "Social-demographic Attitudes and Violence"; Baron, Kennedy, and Forde, "Male Street Youths' Conflict".

need to draw a distinction between *people* and *offenders*. In fact, one possible explanation for the low crime rate of non-socially deprived people is precisely that non-socially deprived people mostly only offend when they faced strong temptations to offend. (Isn't it intuitive that, for example, most rich people only commit crimes when they think they can get away with it while the poor people may take more risks in committing similar crimes?) Without ruling this possible explanation out or without a survey of the mental states of the offenders, I cannot see how we can safely claim that socially deprived offenders on average faced stronger temptations to offend when they offended.

Now let us consider the suggestive argument against the empirical claim of the motivating strength argument, which appeals to upbringing (and which is closely connected to my first objection against the suggestive argument for the claim). If social deprivation correlates with bad upbringing and if bad upbringing correlates with having a diminished concern for moral norms, which I think is plausible, then many socially deprived offenders may in fact have acted for weak temptations in offending whether or not they also happen to face strong temptations to offend when they offended. I do not know how strong this argument is – while it is plausible to believe that social deprivation is somehow connected to concern for moral norms,[93] I do not think we know much about how strong the connection is. (The answer may, of course, be different for different moral norms and different from society to society.) But that precisely explains why I think it is premature to draw any conclusion about the empirical claim of the motivating strength argument at present and we should rather remain agnostic before further empirical evidence is offered.

6. Conclusion

In this chapter I have assessed two arguments for claiming there is a partial excuse that is more applicable to socially deprived offenders. The first argument, the temptations argument, is unsuccessful as the moral claim of the argument is false. The moral claim of the second argument, the motivating strength argument, which emerged from our discussion in section 3, is more plausible. However, I have argued that we do not have sufficient evidence to accept the empirical claim of the motivating strength argument yet, given our present state of knowledge.

---

93 This receives indirect support from the literature on the relationship between social class and their values. See, e.g. Lopez-Calva, "Is There Such Thing As Middle Class Values?"; Argyle, *The Psychology of Social Class,* chapter 10.

Chapter 3: The prior fault argument

1. Introduction

All voluntary conducts, including criminal conducts, are motivated[94] by the mental states (desires, beliefs, values, etc.) of their agents at the time of performing the conducts.[95] Different offenders may of course differ in terms of *what* mental states motivated their crimes. But they may also differ along another dimension, namely, *how they acquired* the mental states that motivated their crimes: were those mental states caused by their past culpable conducts? For example, two thieves who were motivated by the same dishonest motivation when they committed their thefts may differ in terms of how they acquired their dishonest motivation – one of them may have acquired his dishonest motivation culpably but not the other.

The argument to be discussed in this chapter focuses *not* on *what* mental states an offender acted on in committing his crime, but rather on *how* he acquired those mental

[94] The idea of motivation is taken from Duff's *Intention, Agency, and Criminal Liability*, 59: to say that some mental states motivate an act means, roughly, that without those mental states the agent would not have so acted. This account is intended to be very broad, and is consistent with both a causal account and a constitutive account of the relationship between mental states and action.

[95] I use the word "conduct" in a broad manner to include both acts and omissions. This is the standard usage in English criminal law.

states.[96] Let me lay out the main claims of the argument to be discussed (and criticized) in this chapter, *the prior fault argument*, before assessing it:

<u>The Prior Fault Argument</u>

Moral claim: If an offender was not culpable for acquiring the mental states that motivated his crime, i.e. those mental states were not caused by his *prior* faults, then he is less blameworthy for that crime and hence there is a reason to punish him less for that crime than an otherwise similar offender who was culpable for acquiring the mental states that motivated his crime.[97] (Let us call the moral claim of the prior fault argument "*the prior fault principle*".)

(I will call the crimes, the punishment of which is at issue in this chapter, "present" crimes to mark clearly the temporal distinction between them and the *past* conducts causing the mental states that motivated those crimes.)

Empirical claim: The proportion of socially deprived offenders who were not

[96] This kind of argument is criticized in Morse, "Severe Environmental Deprivation", 160, 163, but Morse focuses on attacking the empirical claim while I will focus on attacking the moral claim.
[97] Some people believe that in order for the moral claim of the education argument to be true, the offender who acquired the mental states that motivated his present crime because of his prior culpable conducts must, in addition, *foresee a risk of committing the present crime in the future when he performed the prior culpable conducts*. See Smith, "Culpable Ignorance", 551; Vargas, "The Trouble with Tracing". I do not care whether this condition is added: I believe the moral claim of the education argument is false even if we add this requirement and my objection against the moral claim of the education argument would not be weakened if we add this requirement to that claim.

culpable for acquiring the mental states that motivated their crimes is greater than that for those who are non-socially deprived offenders.

Before assessing the prior fault argument, I wish to make four clarifying remarks about the argument.

First, the terms “culpable” and “fault” can, in ordinary language, include the *kinds* of fault that are *not* relevant to criminal punishment. For example, suppose I was forgetful and as a result I acquired a certain mistaken belief. One may say that I was culpable or at fault for acquiring the belief. After all, forgetfulness is a kind of defect and if my forgetfulness leads to damage of your property I may bear civil responsibility for that. But given that forgetfulness does not seem to be a kind of defect that *criminal punishment* should be concerned with,[98] it is implausible to believe that there is a reason to impose a harsher sentence for my crime simply because I acquired the beliefs that motivated my crime as a result of my forgetfulness in the past. As such, in order for the prior fault argument to have some *initial plausibility*, the ideas of “culpable” acquisition of mental states and prior “fault” have to be understood as referring, *not* to any kind of fault or culpability, but to *the kinds of*

[98] Tadros, *Criminal Responsibility*, 250-254.

culpability and fault that are relevant to criminal punishment. I will not defend a particular position as to what kinds of fault are relevant to criminal punishment here. The point I am making is simply that no matter how one draws the distinction between the kinds of fault that are relevant to criminal punishment and other kinds of fault, the prior fault argument has to be formulated with that distinction in mind.

Second, the prior fault argument is a *historical* argument. Some mitigating arguments appeal to the *present features* of offenders. By a "present feature" I mean a description about the offender the truth of which depends on *the time when he committed his crime*. The temptations argument and the motivating strength argument, which I discussed in the last chapter, are non-historical arguments. For example, the temptations argument claims that there is a reason to punish an offender who faced stronger temptations to offend *at the time of offence* less for his crime. The history of offenders only has a *non-intrinsic* significance under those arguments: it *may* be the case that offenders who had a certain kind of history are more likely to face stronger temptations at the time of offence, but *ultimately* what matters under the temptations argument is whether the offender faced stronger temptations when he committed his crime, not his history. In contrast, the prior fault argument does *not* focus on the *present features* of offenders but rather on their *historical features*. By a historical feature I mean a description about the offender the truth of which depends on *his history before the occurrence of the crime*.[99]

To illustrate the historical nature of the prior fault argument, consider two adult gangsters who attacked members of the opposing gang under the belief that one should, morally speaking, obey the orders of one's gang. They were *identical* except that one of the gangsters had a middle class upbringing when he was young whereas the other gangster grew up in slums. Since they were *identical at the time of the assault*, no arguments that rely on the present features of offenders can justify imposing a different punishment on them for their respective assault. However, if it can be shown that the gangster with a middle class upbringing was culpable for associating with gangsters when he was young, which led to his eventual acceptance of the gangsters' moral code, but that is not the case for the gangster who grew up in slums, then, according to the prior fault argument, there is a reason to punish the latter gangster less for the assault.

Third, while the prior fault argument, as a historical argument, is distinct from arguments that rely on the present features of socially deprived offenders, the two

[99] For clear illustrations of the distinction between historical and non-historical arguments, see Talbert, "Implanted Desires, Self-formation, and Blame", 2-3; Vuoso, "Background, Responsibility, and Excuse".

types of argument are not mutually exclusive.

Fourth, there are *many* conceptually possible historical arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders. The prior fault argument, which I consider in this chapter, is *only one particular version of* historical arguments. Let me illustrate this point by contrasting the prior fault argument with two other types of historical argument. First, recall the distinction I drew in the introductory chapter between exculpatory arguments and non-exculpatory arguments. An exculpatory argument claims that there is a reason to punish some offenders less for their crimes *because* they are less *blameworthy* for those crimes. A non-exculpatory argument, in contrast, claims that there is a reason to punish some offenders less for their crimes *even if* they are *not* less blameworthy for those crimes. The prior fault argument is an *exculpatory* argument and as such is distinct from historical arguments that are non-exculpatory. I will consider a few historical non-exculpatory arguments in parts 3 and 4 of this thesis. Second, that the offender's crime was not caused by his prior faults is only one historical feature amongst many. As such, the prior fault argument is also different from *other exculpatory* historical arguments that rely on *different historical features*. To illustrate, consider the following argument, which can be called *the education argument*. According to the education argument, socially deprived offenders are more likely to have a bad moral education; and if an offender had a bad moral education, then he is less blameworthy for his crime than an otherwise similar offender who had a good moral education and accordingly, there is a reason to punish him less for his crime, *no matter whether (and to what extent) their crimes are a result of their prior faults*. The education argument is a historical argument; it is also an exculpatory argument; but it is different from the prior fault argument as the two arguments appeal to different historical features ("the offender was not culpable for acquiring the mental states which motivated his present crime" and "the offender had a bad moral education" respectively – while empirically these two features are correlated conceptually they are distinct) in justifying the conclusion that there is a mitigating factor that is more applicable to socially deprived offenders. I will consider the education argument in chapter 4.

After clarifying the prior fault argument, let us look at its plausibility. The prior fault argument has some intuitive appeal. Let us start by considering a case outside punishment, which can be called the *student exam case*. Compare two students who both did poorly in their exams. One of them, however, received a very good academic education while the other did not. I understand a good academic education as having both conative and cognitive components: it is easier, and more likely, for a person who had a good academic education to develop the mental states (including the relevant desires and beliefs) that enable one to do well academically. (A good academic education, of course, is not fool-proof: it is possible, though less likely, for a person with a good academic education to fail to develop the relevant mental states.) Many people believe that the student who did not receive a good academic education was *more likely to be not blameworthy for failing to acquire the academic qualities* that would enable him to do well in the exam, like the relevant mathematical beliefs (for the student with a good academic education, in contrast, we are more likely to trace his failure to acquire the correct mathematical beliefs to his prior faults like laziness in preparation); and *if that is the case* then there is a reason to blame him less for his poor performance in the exam and this should be reflected in the associated prizes or reprimands. Analogously, isn't it plausible to believe that socially deprived offenders, who grew up in environments that are conducive to the acquisition of criminogenic mental states, were more likely to be *not blameworthy for acquiring the mental states that motivated their present crimes*; and as a result, they are often less blameworthy for their present crimes; and accordingly there is often a reason to punish a socially deprived offender less for his present crime?[100]

[100] Lippke, "Diminished Opportunities, Diminished Capacities", 471. For this line of thought, see also Fischer and Ravizza, "Introduction", 33.

I believe the prior fault argument is unsuccessful even if the empirical claim of the argument is true, because the prior fault principle is false and I will argue against the prior fault principle in the rest of this chapter. But I wish to explain right at the outset that my position, which involves the rejection of the prior fault principle, is not as radical as it seems. Many people have the intuition that there is a reason to punish an offender who acquired the mental states that motivated his present crime culpably more *in total* than an otherwise similar offender who acquired the mental states that motivated his present crime non-culpably. My rejection of the prior fault principle actually does *not* involve rejecting that intuition.

To see this point clearly, we need to remind ourselves that our topic is how much punishment an offender should receive *for his present crime, i.e. how much punishment he should receive for the conviction based on his present crime*. Since a person can be convicted of multiple crimes, we need to draw a distinction between the amount of punishment imposed on an offender *for his present crime* and the amount of punishment imposed on him *in total*. The prior fault principle is a principle about how much punishment a person should receive *for his present crime.* In rejecting the prior fault principle, I argue that there is no reason to punish an offender more *for his present crime just because the mental states that motivated that crime were caused by* *his prior faults*. I will *not*, however, deny that there is a reason to punish him more *in total* as I believe that there is a reason to impose separate convictions and punishments for his prior faults.

To illustrate the exact way in which I disagree with the principle of prior fault, let us consider the following case, which I will call the *addicted thieves case*. Imagine two adult drug addicts who just stole something in order to satisfy their addiction. (I am assuming that while the addiction provided some extra desires at the time of theft, the desires were not irresistible or so strong such that it is unreasonable to expect the agents to resist them, i.e. the agents were not compulsive or involuntary when they were stealing.) The only difference between them is that one of the addicts culpably experimented with drugs with knowledge of the risks when he was young, while the other addict acquired the addiction involuntarily when he was young (say, because someone injected the drug into his body against his will at that time). I will argue, contrary to the principle of prior fault, that there is no reason to punish the culpably addicted thief more for *theft*. However, I do not deny that there is a reason to punish him more *in total* as I believe that there is a reason to impose a separate conviction and punishment on the culpably addicted thief *for his prior fault*, i.e. *trying drugs* *culpably*.[101]

The rest of this chapter will be arranged as follows. In section 2, I will reply to an objection that claims that, given my clarification above as to the exact way in which I disagree with the prior fault principle, my dispute with the prior fault principle is *trivial*. In section 3, I will argue against the prior fault principle. My case relies heavily on expressive considerations.

In section 4, I will revisit the student exam case (which I mentioned above), a case outside punishment. Although prizes and reprimands are different from criminal punishment, they are also expressive and hence I believe the arguments I raise in section 3, with suitable modifications, also apply to them. I will suggest how the

[101] Rebecca Williams adopted a related proposal in relation to intoxication cases. Suppose a person attacked someone under a state of automatism caused by reckless intoxication. She argues that the intoxicated agent should not be convicted of assault but he should be convicted of a separate offence of "committing the actus reus of assault while recklessly intoxicated". See her "Voluntary Intoxication – a Lost Cause?".

I have learnt much from reading her paper. There are, however, important differences between her proposal and mine. First, her proposal covers only cases where the defendant lacked mens rea (or voluntariness) at the time of committing his present conduct, and thus would not apply to the cases that I am focusing on like the gangsters case and the addicted thieves cases where the agents had the mens rea for an offence at the time when they committed their present conducts. Second, her proposal only covers cases where the prior fault is one of reckless intoxication, while my proposal is meant to have a broader application. Third and related to the first difference between our proposals, I believe my defence of my proposal is significantly different from her defence of her proposal. The main reason Williams offered for convicting the recklessly intoxicated agent of the separate offence is that it would be unfair to convict him of the offence based on his present conduct (e.g. assault), given that he did not have the mens rea for that offence. See Williams, "Voluntary intoxication – a Lost Cause?", 275-276. That rationale, however, does not apply to my case since there is no unfairness in, for example, convicting the culpably addicted thief of theft (since it is stipulated that he had mens rea for theft at the time of stealing) and I am not arguing that he should not be convicted of theft. I appealed to different considerations to justify imposing a separate conviction on him in section 3.3.2 below.

student exam case should be properly dealt with by applying those arguments.

2. The triviality objection

I have tried to avoid the charge of counter-intuitiveness by claiming that my dispute with the prior fault principle does *not* involve rejecting the intuitive idea that there is a reason to punish an offender who culpably acquired the mental states that motivated his present crime more *in total* than an otherwise similar offender who committed the same present crime but who acquired the mental sates that motivated the crime non-culpably – I only argue that there is no reason to punish the former more *for his present crime*. But in avoiding the charge of counter-intuitiveness, my opponent may claim that my position falls prey to the charge of *triviality*. This is because, according to this objection, which can be called *the triviality objection*, the *only* thing we care about is how much *total* punishment a person receives and we do not care *for what* the punishment is imposed. Using the addicted thieves case above to illustrate, according to my opponent, *once* we agree that there is a reason to punish the culpably addicted thief more *in total* than the involuntarily addicted thief, we *need not care* whether there is a reason to punish him more *for theft*, or rather that there is no reason to punish him more for theft while there is a reason to punish him more *for trying drugs* *in the past*.

Putting this point back in the context of my thesis, according to the triviality objection, the *only important issue* is whether there is a reason to punish a non-socially deprived offender more *in total* and it is *unimportant* whether there is a reason to punish him more in total because there is a reason to punish him more *for his present crime* or rather because there is a reason to punish him more for his *prior faults.* Since my critique of the prior fault argument, even if sound, does *not* reject the claim that there is a reason to punish a non-socially deprived offender more *in total*, it is not, according to the triviality objection, a substantial critique of the prior fault argument.

I think the triviality objection is unsound as we care *not only* about the total amount of punishment a person receives, but also about whether the convictions and the amount of punishment imposed for each conviction are proper. To see this point, consider two cases. The first case can be called the *different crime case*. Suppose that a person is convicted of a murder that happened yesterday and life imprisonment is imposed for that. The defendant actually did not kill anyone yesterday – the police got the wrong person. Unbeknownst to others, he committed a totally unrelated brutal rape twenty years ago. In such a case, in imposing life imprisonment on the defendant for the murder that happened yesterday, we need not impose on him a greater total amount of punishment than what he deserves (assuming that life imprisonment is the pro*per se*ntence for both murder and brutal rape). But still *some* injustice is clearly done.[102] The second case can be called the *quantum exchange case*. Suppose a person committed murder and petty theft independently and he has been duly convicted of them. The court then imposed a very heavy punishment for petty theft and an insignificant punishment for murder. The total amount of punishment imposed on the offender may be correct. But again it seems that some injustice is done.

One reason (I am not sure whether it is the only reason) why we care, not only about the correctness of the total amount of punishment imposed, but also about whether the convictions and the amount of punishment imposed for each conviction are proper, is because of the *expressive* aspect of criminal punishment. Criminal punishment is by its nature expressive.[103] Once we take into account the expressive nature of criminal punishment we can understand why there is an injustice in the different crime case and the quantum exchange case: the first offender is wrongfully condemned for a wrong that he has not committed; the second offender is censured for his petty theft in a way that does not correspond to his blameworthiness for that crime.

---

[102] This case is not my original creation. I remember reading it somewhere but I am afraid that my reference on this case is mistaken so I cannot trace the actual source at present.

[103] Feinberg, "The Expressive Function of Punishment"; von Hirsch, *Censure and Sanctions*.

3. My critique of the prior fault principle

3.1. Distinction between an agent's direct and indirect blameworthiness for a conduct

As long as we are clear that we are talking about the blameworthiness for the present crime but not the prior faults leading to the present crime, it may be claimed that it is *obvious* that the prior fault principle is mistaken. In the different crime case, while the offender may be blameworthy for the brutal rape twenty years ago, he is not blameworthy for yesterday's murder. In the quantum exchange case, while the offender may be very blameworthy for murder, he is not very blameworthy for petty theft. Analogously, is it not obvious that while the offender who acquired the mental states that motivated his present crime because of his prior faults may be more blameworthy *for his prior faults,* he is not more blameworthy *for his present crime*, and accordingly, there is no reason to punish him more *for his present crime* based on proportionality?

I believe this objection against the prior fault principle is too quick. In order to understand why such a simple disposal of the prior fault principle is too hasty, let us, following many writers on moral responsibility, draw a distinction between *direct* (or original) and *indirect* (or derivative) blameworthiness.[104] A person's indirect blameworthiness for a present crime depends on how much the present crime is a *consequence* of other past conducts for which he is directly blameworthy and how directly blameworthy he is for those past conducts. In contrast, a person's direct blameworthiness for a present crime is not derived from his blameworthiness for other conducts.

I believe that the offender who acquired his mental states that motivated his present crime culpably is *not* more *directly* blameworthy for his present crime. This is because I believe that a person's direct moral blameworthiness for a wrongful conduct depends only on his quality of will,[105] i.e. whether he has sufficient concern for others, in performing the conduct, and a person's quality of will in acting, while depending on

[104] E.g. Smith, "Culpable Ignorance"; Adams, "Involuntary Sins"; Rosen, "Culpability and Ignorance"; Rosen, "Skepticism about Moral Responsibility", 298-299.

[105] There are two possible exceptions, none of which affects my main point here. First, wrongfulness may affect blameworthiness without affecting quality of will. But clearly this is irrelevant here since whether one acquired the mental states that motivated his act because of his prior faults is irrelevant to the wrongfulness of the act. Second, a person's direct blameworthiness for a conduct may depend, in addition to his quality of will when he performed the conduct, the psychological connectedness between that conduct and the person whom we are trying to hold responsible. I will discuss the psychological connectedness requirement in chapter 4. However, we should observe that this exception does not provide any assistance to the prior fault argument. This is because the issue of whether the person whom we are trying to hold responsible is psychologically connected to a conduct depends on what happened *after* the conduct (say, whether he reformed after the conduct and changed into another person when we try to hold him responsible) but not on what happened *before* the conduct like whether he acquired the mental states that motivated the conduct culpably.

his mental states when he acts, does *not* depend on *how* he acquired those mental states. I will defend the quality of will view of direct moral blameworthiness in chapter 4 of this thesis below.

But I think even if my opponent concedes this point about direct blameworthiness to me, and accordingly agrees with me that the offender who acquired his mental states that motivated his present crime culpably is not more directly blameworthy for his present crime, there is still *a clear sense* in which that offender is more blameworthy for his present crime, namely, he is more *indirectly* blameworthy for his present crime. For example, while I believe that the culpably addicted thief is *not more directly blameworthy for theft* than the involuntarily addicted thief, I agree that he is more *indirectly blameworthy for theft* since his theft is the consequence of a past conduct for which he is directly blameworthy (i.e. trying drugs).

Accordingly, once we draw a distinction between direct and indirect blameworthiness, we can see why we cannot dispose of the prior fault principle so quickly. In the different crime case, the offender is not blameworthy for yesterday's murder *in any sense*; similarly, in the quantum exchange case, the offender is not very blameworthy for the petty theft *in any sense*. In contrast, there is *a sense* in which the offender who acquired the mental states that motivated his present crime culpably is more blameworthy for his present crime, namely, he is more *indirectly* blameworthy for his present crime.

But while the principle of proportionality, in itself, does not specify whether *the kind* of blameworthiness that ought to correspond to the punishment for a crime is the agent's direct or rather indirect blameworthiness for that crime, I will argue, in the rest of section 3, that *nonetheless* there are strong *expressive* grounds to make an offender's punishment for a crime proportionate to his direct blameworthiness for the crime only *without regard to* his indirect blameworthiness for the crime. I will argue, in section 3.2, that *if* separate convictions and punishments for the prior faults have been imposed on an offender who is indirectly blameworthy for his present crime, *then* there is no reason to take his indirect blameworthiness for the present crime into account in deciding the punishment for the present crime; and I will argue in section 3.3 that, assuming that some prior faults, which led to a present crime, should be taken into account in criminal punishment, there are strong expressive grounds for imposing separate convictions and punishments for those prior faults instead of taking them into account in deciding the punishment for the present crime.

3.2. If separate convictions and punishments for the prior faults have been imposed on an offender who is indirectly blameworthy for his present crime, then there is no reason to take his indirect blameworthiness for the present crime into account in deciding the punishment for the present crime

I think most people, even the proponents of the prior fault principle, would agree that *if* separate convictions and punishments for the prior faults have been imposed on an offender who is indirectly blameworthy for his present crime, *then* there is no reason to take his indirect blameworthiness for his present crime into account in deciding the punishment for the present crime. To illustrate with the addicted thieves case, suppose the proponent of the prior fault principle originally suggested that the culpably addicted thief should be punished for x units more for theft than the involuntarily addicted thief. But suppose we have now independently convicted the culpably addicted thief *of drug taking* and we have already imposed an x-unit punishment on him for that. *Once that has been done*, it seems that even the person who originally suggested that the culpably addicted thief should be punished x units more for theft would retract his original suggestion. Otherwise he would seem to be *double-counting* the culpably addicted thief's culpability for trying drugs, *first* taking that into account in deciding the punishment *for drug taking*, *then* taking that into account *once again* in deciding the punishment *for theft*. In other words, I think even the proponents of the prior fault principle would regard the principle to be a *conditional* one, as applying *only where* separate convictions and punishments have not been imposed for the prior faults causing the present crime.

3.3. If a prior fault, which caused a present crime, should be taken into account in criminal punishment, then it should be taken into account separately from the punishment for the present crime

In section 3.2, I argued that *if* separate convictions and punishments for the prior faults have been imposed on an offender who is indirectly blameworthy for his present crime, *then* there is no reason to punish him more for his present crime than an otherwise similar offender who is not indirectly blameworthy for his present crime. In other words, I argued that *if* a prior fault leading to the present crime has been taken into account in criminal punishment separately, *then* there is no reason to take it into account *again* in the sentencing for the present crime.

But I have not yet examined the issue of, assuming that a prior fault that caused a present crime should be taken into account in punishment (i.e. the offender whose present crime was traceable to the prior fault should be punished more *in total*), which one of the following two proposals we should adopt:

(1) Imposing a separate conviction and punishment for the prior fault and *not* taking the prior fault into account in deciding the punishment for the present crime caused by it; or rather

(2) Taking the prior fault into account in deciding the punishment for the present crime caused by it.

The difference between my proposal (i.e. proposal (1)) and my opponent's proposal (i.e. proposal (2)) can be put into sharper focus if we consider the addicted thieves case again. Assuming that the prior fault of the culpably addicted thief should be taken into account in criminal punishment (i.e. he should be punished more *in total* than the involuntarily addicted thief), how will he be treated under the two proposals? Under my opponent's proposal, he will be convicted of *only one offence* (theft) and receive *one single sentence*. The fact that the motivating mental states of theft were caused by his prior fault would be regarded as *a reason for giving him a more severe sentence for theft than the involuntarily addicted thief.* (For the purpose of illustration, let us stipulate that the culpably addicted thief receives a three-year sentence for theft while the involuntarily addicted thief receives a two-year sentence for theft under my opponent's proposal.) In contrast, under my proposal, he will be convicted of *two offences* (theft *and* trying drugs) and receive two separate sentences (his sentence for theft would be the same as the involuntarily addicted thief and the sum of his two sentences would be equal to the sentence for theft under my opponent's proposal). (Using the example above to illustrate, the culpably addicted thief will receive a two-year sentence *for theft* and a one-year sentence *for trying drugs*, while the involuntarily addicted thief will receive a two-year sentence for theft, under my proposal.) At first glance, it may seem that there is not much basis to choose between the two proposals. I disagree. I will argue for my proposal in the rest of section 3.3.

The rest of section 3.3 is divided into three parts. In section 3.3.1, I will explain why I am *not* committed to the idea that all prior faults should be taken into account in punishment all things considered or a particular theory of what kinds of prior fault should be taken into account in punishment all things considered. In section 3.3.2, I will state the three expressive arguments for preferring my proposal over my opponent's. In section 3.3.3, I will consider some objections against my proposal.

3.3.1. That I am not committed to the idea that all prior faults should be taken

into account in punishment all things considered or a particular theory of what kinds of prior fault should be taken into account in punishment all things considered

Before I go on, let me stress that in arguing for my proposal over my opponent's, I am *not* committed to the view that *all* prior faults of an offender that led to his present crime should be taken into account in punishment, all things considered. This should be clear once we pay a closer attention to the dialectical structure.

My aim is to argue against the prior fault principle, the claim that there is a reason to take prior faults that caused a present crime into account in punishment *through affecting the sentence for the present crime*. My main point is that *if* some prior faults, which caused a present crime, should be taken into account in punishment, all things considered, *then* we should prefer doing so separately from the present crime (i.e. imposing separate convictions and punishments for them) *over* taking them into account in deciding the punishment for the present crime. I am entirely open to the suggestion that some prior faults (even those that concern the kinds of fault relevant to criminal law – see section 1 of this chapter above) should not be taken into account in criminal punishment at all, all things considered. Perhaps due to pragmatic or political reasons we should not investigate very deeply into an offender's past.[106] That may or may not be true but that would not affect my criticism of the prior fault principle.

A related point is that I am *not* arguing for a particular theory of *what* kinds of prior fault should be taken into account in criminal punishment all things considered. Instead, my point is that *whatever kinds of* prior fault a proponent of the prior fault principle believes ought to be taken into account in criminal punishment all things considered (All prior faults? Prior faults in the last ten years? Prior faults that the state can verify without too much intrusion of privacy? Prior faults that a sentencing system can take into account without too much administrative cost?), we should prefer taking them into account separately from the present crime instead of taking them into account in deciding the punishment for the present crime.

3.3.2. Three expressive arguments for my proposal

Now let me lay out the three expressive arguments for favouring my proposal over my opponent's. The first argument can be called the *fair labelling of the prior fault argument*. Other things being equal, we want to impose convictions on an offender in

[106] Ashworth, "Justifying the Grounds of Mitigation".

a way that provides more information about the wrongfulness of his conducts.[107] That explains why it is inappropriate to abolish the offence of murder and convict murderers of common assault even if we impose suitably heavy sentences on this particular group of "assaulters": to convict murderers of common assault alongside with attackers who simply touched their victims without causing harm would not mark the wrongfulness of murders sufficiently.[108] I think my proposal provides a fairer labelling of the prior faults than my opponent's proposal in many cases and this can be illustrated when we think about the culpably addicted thief again. Under my opponent's proposal, the culpably addicted thief will only be convicted of theft. While his sentence will be longer than the involuntarily addicted thief because of his past culpable drug consumption, his prior fault would not be marked in his convictions and criminal record. In contrast, under my proposal, he will be convicted, in addition to theft, of the offence of trying drugs. In such a case, the wrongfulness of his prior fault is more adequately described in his convictions and criminal record.

The second argument can be called the *clear marking of the time of prior fault argument*. Under my proposal, the offender will be convicted of his prior faults separately from his present crime. Since the criminal record of a person usually marks

[107] For discussion of this idea, see Leverick and Chalmers, "Fair Labelling in Criminal Law".
[108] Ashworth, "The Elasticity of Mens Rea", 53.

the dates when the convicted conducts were committed, under my proposal it is clearer to the public when the prior faults occurred. In contrast, under my opponent's proposal, the offenders will not be convicted separately of their prior faults; the prior faults will only be taken into account as sentencing factors for the present crime and one's criminal record does not usually contain details about sentencing factors.

One may ask why it is important to know *when* the prior faults occurred. The reason is that, for many practical purposes, we have good reasons to discount a person's culpable conduct depending on the time lapsed since the conduct (this explains why many criminal justice systems consider certain sentences as spent after a certain period of time).[109] A person who assaulted someone yesterday may be quite dangerous, but that is less plausible if the assault happened twenty years ago. Similarly, we may have different views towards an offender who committed his prior fault yesterday from someone who committed his prior fault twenty years ago.

The third argument can be called the *clarity of direct blameworthiness argument*. It is valuable for the court to send out an unambiguous message about an offender's direct blameworthiness for his present crime. My opponent's proposal would send out a

[109] E.g. for England, see *The Rehabilitation of Offenders Act 1974*.

more ambiguous message about the offender's direct blameworthiness for his present crime. To illustrate this point, suppose a person receives a six-year sentence for rape. Under my opponent's proposal, it is not clear, without further clarification, whether the six-year sentence reflects his direct blameworthiness for rape, his indirect blameworthiness for rape, or a combination of both. This ambiguity is undesirable partly because a total amount of punishment justified by multiple separate lesser wrongs may reflect very differently on the offender from the same amount of punishment justified by a single serious wrong.[110] This difference in attitude is recognised in the English criminal justice system. For example, if an offender was convicted on three different occasions and he received a two-year imprisonment for each of those convictions, the sentences will all be regarded as spent after a certain period of time; but a sentence of six-year imprisonment for a single crime will never be regarded as spent.[111] A closely related problem is that under my opponent's proposal, it is possible for a type of crime that is less serious to receive generally heavier sentences. For example, suppose that rapists did not often acquire the mental states that motivate their rapes culpably, but assaulters often acquired the mental states that motivate their assaults culpably. If that is the case, then if we take prior faults that led to a present crime into account in deciding the sentence for the present

[110] For a similar idea, see Mitchell, "Multiple Wrongdoing and Offence Structure", 403; Horder, "Rethinking Non-fatal Offences against the Person", 348.

[111] *The Rehabilitation of Offenders Act 1974*.

crime, then it is possible for sentences for assault, which I take to be a less serious crime than rape, to be generally heavier than sentences for rape. This may mislead the public as to the state's view on the relative seriousness of rape and assault.[112]

Let me make two qualifications about the three arguments that I have just offered for favouring my proposal over my opponent's. First, I do *not* wish to suggest that they are *non-contingent* arguments that apply to all possible societies. The soundness of the three arguments depends on certain empirical assumptions about our society. Let me highlight two. First, the fair labelling of the prior fault argument and the clear marking of the time of prior fault argument depend on the assumption that information about convictions is more readily available than the sentencing judgments explaining the grounds for the sentence: otherwise there would not be much extra value in labelling the offenders' prior faults and recording the dates of the prior faults clearly through convictions as these information would, after all, be available in the sentencing judgments. Second, the clarity of direct blameworthiness argument assumes that information about the grounds for a sentence is less readily available than the total gravity of that sentence. If information about the grounds for a sentence is as readily available as the total gravity of that sentence, then people can easily find out how

[112] For a related point, see von Hirsch and Ashworth, *Proportionate Sentencing*, 70-71.

much of a sentence for a crime is for the agent's direct blameworthiness for the crime and how much of the sentence is for the agent's indirect blameworthiness for the crime even under my opponent's proposal. That said, I believe these empirical assumptions are true for many Western societies. Most members of the public do not (with good reasons based on efficiency) enquire very deeply into the sentencing judgments and grounds for the sentences. Criminal records, which most potential employers and people dealing with the offenders look at, do not contain the grounds for the sentences either.[113]

Second, I do not wish to suggest that expressive concerns are the most important concerns in sentencing. (In fact, while I believe that expressive concerns in punishment are important, I will argue in chapter 8 of this thesis that they are not the lexically trumping concerns in sentencing.) Rather, I have a more modest aim in employing the three arguments based on expressive concerns: I believe my proposal does not fare worse in other dimensions, and given that it fares better in terms of expressive concerns, it should be favoured. Whether I am right in believing that, of course, depends on whether my proposal faces otherwise strong objections. I will now consider five objections against my proposal. Replying to these objections will also

[113] Leverick and Chalmers, "Fair Labelling in Criminal Law", 223, 234; Ashworth, "The Elasticity of Mens Rea", 56.

enable me to clarify *how* exactly I think the prior faults should be criminalised.

3.3.3. Objections against my proposal

The first objection against my proposal, the *net of liability objection*, claims that in many cases we cannot impose a separate conviction and punishment for the prior fault *as we would cast the net of criminal liability too wide if we criminalise the prior fault separately*. To illustrate this objection, consider an adult attacker who acquired his mental states that motivated his present assault as a result of his culpable association with gangsters when he was young. According to the net of liability objection, his prior fault, i.e. his culpable association with gangsters, should be taken into account in criminal punishment. That is to say, this attacker should be punished more *in total* than an otherwise similar attacker who acquired the mental states that motivated his assault non-culpably. This objection continues by claiming that we *cannot*, however, as my proposal would suggest, impose a *separate* punishment on him for his prior fault (i.e. culpable association with gangsters). *This is because we do not want to cast the net of criminal liability so widely as to impose criminal liability on a person just as soon as he culpably associates with gangsters,* i.e. we do not want to convict and punish a person if he had culpably associated with gangsters *but did not do anything* *bad because of such association afterwards*, and accordingly *we should not separately criminalise culpable association with gangsters*. Therefore, in this case, the *only* way to take into account his prior fault (culpable association with gangsters) that led to his present crime (assault) in criminal punishment is to punish him more for his present crime. The objection then continues by claiming that this point can be *readily generalised*: there are *many* culpable conducts that should be taken into account in criminal punishment *in case* they cause a substantive crime but we do not wish to impose criminal liability on a person as soon as he engages in those conducts.[114] (By "substantive crimes" I mean crimes *other than those that I propose to create to cover the prior faults*, i.e. the kinds of crime that we have at present. Examples of "substantive crimes" are assault, theft, murder, rape, attempted murder, etc.)

The net of liability objection presupposes that there are some culpable conducts (including but not limited to culpable association with gangsters) that, on the one hand, should be taken into account in criminal punishment *in case* they cause a substantive crime like assault; while on the other hand, a person should *not* attract criminal liability just *as soon as* he engages in these conducts. This invites a question: if a

[114] A similar objection has been raised against the proposal that we should abolish the Majewski rule and simply criminalise reckless intoxication *per se*. The objection is that to do so would make liability too wide, as reckless intoxication, *per se*, should not be made into an offence. See Williams, "Voluntary Intoxication: a Lost Cause?", 276.

culpable conduct, like culpable association with gangsters, should be taken into account in criminal punishment when it caused a substantive crime like assault, then why should we not also impose criminal liability for that conduct when it does not cause further bad consequences? Two answers can be given by a proponent of the net of liability objection. The first relies on cost concerns. If we, for example, impose criminal liability on people as soon as they culpably associate with gangsters (i.e. even if such association with gangsters does not cause any further bad consequences) then we will send many more people through the criminal justice system, which is costly. The second relies on the more controversial idea of moral outcome luck. Many people believe that a person's blameworthiness for a culpable conduct depends on the outcome. For example, many people believe that a reckless driver who caused death is more morally blameworthy for his reckless driving than another driver who is equally reckless but who luckily did not cause death.[115] If we believe that a conduct, even if culpable, should merit the attention of criminal punishment only if it is *sufficiently* blameworthy, then that provides another reason for why many culpable conducts satisfy the following two criteria: first, they should be taken into account in criminal punishment if they cause a substantive crime like theft or assault, but second, we should not impose criminal liability on a person just as soon as he engages in these

[115] Scanlon, *Moral Dimensions*, 126.

conducts.

In reply to the net of liability objection, I agree that there are many culpable conducts, including culpable association with gangsters, that on the one hand, should be taken into account in criminal punishment if they cause a substantive crime like assault or theft, while on the other hand, a person should not attract criminal liability as soon as he engages in these conducts. *But I do not believe it follows that these conducts should not be criminalised*, and accordingly, I do not believe that when these conducts caused a present crime like assault, we cannot punish these conducts separately from the present crime. My reply, which draws on the work of Williams, relies on a distinction between two ways of criminalising a conduct.[116] First, we can make a conduct an offence *regardless of the consequences of the conduct*. Rape is an example of such an offence: a person commits rape as long as he engages in non-consensual sex irrespective of the further consequences. Second, we can create a *conditional* offence, i.e. make the conduct an offence *only if it caused certain specified consequences*. For example, gross negligence is not *per se* a criminal offence under English law; but if death was caused by a person's gross negligence, he is liable for

[116] Rebecca Williams believes that we should not criminalise reckless intoxication *per se*. Yet, she believes that if one's reckless intoxication resulted in committing the actus reus of a substantive crime, say assault, then we can punish him for such intoxication through the conditional offence of "committing the actus reus of assault while being recklessly intoxicated". See Williams, "Voluntary Intoxication: a Lost Cause?", 276.

manslaughter.

I suggest we can criminalise, as *conditional* offences rather than as unconditional offences, conducts that satisfy the following two criteria: (a) they should be taken into account in criminal punishment if they cause a substantive crime like theft or assault but (b) we should not impose criminal liability on a person just as soon as he engages in these conducts. Using the gangster assault case to illustrate, I agree that we should not criminalise culpable association with gangsters *per se*; but we can make "culpable association with gangsters *that caused a substantive crime*" an offence (we should actually, because of the value of fair labelling, make several different conditional offences depending on the type of the substantive offence caused by the prior fault).[117] If this conditional offence is in place, then we can punish the attacker who acquired the mental states that motivated his assault as a result of his culpable association with gangsters when he was young more in total without punishing him more *for his assault* since we can separately impose a punishment on him *for the conditional offence based on his prior fault* (i.e. the offence of "culpable association with gangsters that caused an assault"). Since creating conditional offences like the ones I just mentioned will not widen the net of criminal liability (for example, the

[117] For the idea that different conditional offences, defined in terms of the result caused, should be enacted, see Williams, "Voluntary Intoxication – a Lost Cause?"

conditional offence I just mentioned will not catch someone who culpably associated with gangsters but did not commit a substantive crime because of such association afterwards), my proposal is not vulnerable to the net of liability objection.

The second objection against my proposal, which can be called the *variety of prior faults objection*, starts with the observation that the prior faults that we should take into account in criminal punishment *come in all shapes.* We have mentioned “culpable association with gangsters” and “trying drugs” as two examples. But clearly there are *many other* types of culpable conduct that, in case they cause a substantive crime later, should be taken into account in criminal punishment. Accordingly, even if we have criminalised (perhaps as conditional offences – see my response to the net of liability objection above) “culpable association with gangsters”, “trying drugs”, and a number of other specific conducts, it is unlikely that we would have criminalised all, or even most, of the conducts that should be taken into account in criminal punishment in case they cause a substantive crime later.[118] Since we cannot create an infinite number of offences, in many cases the prior faults leading to a present crime, which should be taken into account in punishment, will not be covered by an existing offence; accordingly, we cannot punish them separately from the present crime and therefore

[118] For some discussion of related issues, see Leverick and Chalmers, “Fair Labelling in Criminal Law”, 239.

the *only* way to take them into account in punishment is to punish an offender whose present crime was caused by such prior faults more *for his present crime*.

In reply, I agree that the prior faults that should be taken into account in criminal punishment come in all shapes and are unlikely to be fully covered by a number of specific offences, i.e. offences that cover particular kinds of conduct. But that *per se* is not fatal to my position. We can, in addition to making specific offences (like the culpable association with gangsters offence), make *general catch-all conditional offences*. For example, we can create offences like "*culpable conduct* that led to assault" and "*culpable conduct* that led to theft". With these general catch-all conditional offences we can cover all the prior faults that we want to take into account in criminal punishment, and accordingly, punish them all separately from the present crimes caused by them. I wish to add that these catch-all conditional offences should be made *in addition to,* rather than replace, the more specific offences covering the prior faults. For example, even if we have a general offence of "culpable conduct that led to theft", I think it is still worthwhile to have a specific offence of "trying drugs that led to theft" and the latter should be the *preferred* charge if a conduct is covered by both offences.[119] (As an analogy, in English law the general offence of criminal

[119] For the idea that a general catch-all offence is necessary but (because of fair labelling) that does not mean that we should discard specific offences, see Horder, "Rethinking Non-fatal Offences against the

damage exists alongside rather than replace the more specific offence of arson.[120])
This is because charging an offender with a specific offence would fare better in terms of fair labelling – the general catch-all offence should be used as a *last resort* to cover the prior faults not otherwise covered by the specific offences.

My reply to the variety of prior faults objection, which suggests that we need to create general catch-all conditional offences like "culpable conduct that led to assault", which are rather unspecific as to the kind of conduct prohibited, may prompt a third objection, which can be called the *legality objection*. According to this objection, legality requires that *offences* be clearly defined and offences like "culpable conduct that led to assault" are too vague to satisfy the legality requirement since "culpable conduct" is such a broad idea. In contrast, according to this objection, the legality requirement applies with a much lesser force in relation to *sentencing*, so it is acceptable to take into account prior faults in the sentencing of the present crime caused by them even if we have not previously defined clearly what are the kinds of prior fault that would be taken into account in sentencing.

In reply, it is important to observe that the general catch-all offences, which I propose

Person", 347.

[120] See *Criminal Damage Act 1971* s 1; Simester and von Hirsch, *Crimes, Harms, and Wrongs*, 204.

to create, like “culpable conduct that led to assault” or “culpable conduct that led to theft”, are *conditional* offences that will trigger *only if* the offender commits a substantive offence (assault, theft, etc.) later. I agree that legality requires more as to the clarity of the conduct elements of substantive offences than the clarity of sentencing factors. But I do not agree that legality requires more as to the clarity of the conduct elements of a *conditional offence that will only trigger when the offender commits a substantive offence later* than the clarity of sentencing factors.

The main reason why legality demands more in relation to the clarity of the conduct elements of substantive offences than the clarity of sentencing factors is that, because of the importance of *liberty*, it is of great importance that a person can avoid *any* involvement with the criminal justice system at all if he chooses so: that is why we think it is more important to make sure that no defendant incurred a conviction that carries a one-day sentence without fair warning of the offence elements than making sure that no offender incurred an *extra* one-day sentence over his existing sentence without fair warning of the sentencing factors. But this concern does not apply to the general conditional offences that I propose to create since even if these conditional offences are created and their conduct elements are defined in a very vague manner (e.g. “culpable conduct”), one can avoid criminal liability altogether simply by not committing any substantive crimes (e.g. theft, assault), which themselves have to be clearly defined in order to satisfy the legality requirement. No matter under my opponent's proposal or under my proposal, one has a clear, and equally clear, way of avoiding contact with the criminal justice system: avoid substantive offences like assault or theft. In other words, the importance of liberty requires that one has clear warning before one crossed the line of criminal liability. But liberty does not provide any reason to decide whether to convict an offender of one offence or rather two offences after he has crossed the line, and hence, does not provide any reason why vague sentencing factors, which will come into play only after one has otherwise stepped across the line of criminal liability, are preferable to the vague conditional offences that I propose to create, which will *also trigger only* after a person has otherwise stepped across the line of criminal liability.

According to the fourth objection, which can be called *the difficulty of providing details of prior faults* objection, there are some cases where the prosecution can prove that an offender's mental states that motivated his present crime are traceable to his prior faults but the prosecution is unable to provide the *details* of those prior faults. For example, consider the culpably addicted thief again. The prosecution might be able to prove that his addiction is traceable to his culpable conducts of drug consumption when he was young. But it would be rather difficult to prove *when exactly* those conducts occurred: did he take drugs on 1 August 1980 or 2 August 1980? How about 4 August 1980? And so on. This objection continues by claiming that if we are not sure about the exact details of some culpable conducts, then we cannot convict the offender based on them; and accordingly, if we want to take those prior faults into account in criminal punishment, we can only take them into account in the sentencing of the present crime caused by them since sentencing, unlike convictions, should allow for a more broad-brushed assessment of the defendant's culpability.

The crux of this objection is that many prior faults can only be taken into account in criminal punishment using a broad-brushed approach to culpability and a broad-brushed approach can be adopted at the sentencing stage but not the conviction stage. Two points can be made in reply. First, some may wonder whether it is a good idea to adopt a broad-brushed approach in sentencing: should the prosecution's task be made that easy at the sentencing stage? Should the prosecution be allowed to rely on some facts in sentencing without proving their details like when exactly they occurred? But I wish to set that point aside. My second, and main, problem with the objection is that I cannot see why we cannot adopt a broad-brushed approach to culpability at the conviction stage. It seems to me that whether a broad-brushed approach or a narrow approach should be adopted at the conviction stage *simply depends on the shape and form of offences we create*. I want to draw a distinction between two kinds of offence. Some offences are single-conduct offences, covering a single conduct at a particular time. Assault is an example of a single-conduct offence. Other offences are multiple-conduct broad-brushed offences, which cover an agent's conducts through a period of time. One example is the offence of "living on earnings of prostitution of others" under Hong Kong law.[121] Once we realise that we can create offences that concern, not a particular specified conduct, but an agent's conducts *over a period of time,* it seems that my proposal can cover offenders like the addicted thief I mentioned in the last paragraph. To illustrate, we can create, in addition to single-conduct offences like "culpable consumption of drug", multiple-conduct broad-brushed offences like "culpable use of drugs through a period of time"; and we can charge the addicted thief, in addition to theft, with the latter offence, which the prosecution may be able to prove without providing details of each individual instance of culpable drug consumption during his youth.

I wish to observe that relying on broad-brushed offences, rather than single-conduct

[121] *Crimes Ordinance* s 137 (Hong Kong).

offences, in covering some prior faults is not without costs. For one thing, it will not mark the time of occurrence of each individual prior fault on an offender's criminal record as clearly as covering them by single-conduct offences. But punishing an offender, who is indirectly blameworthy for his present crime, for his prior faults separately from the present crime using these broad-brushed offences *at least* marks the broad temporal dimension of the prior faults, while under my opponent's proposal the time of occurrence of the prior faults will *not be marked at all.*

The fifth objection, which can be called the *overburdening the jury objection*, claims that my proposal, which involves imposing separate convictions and punishments for the prior faults instead of taking them into account in the sentencing of the present crime, would overburden the jury. This objection claims that for some offenders, their present crimes were caused by a large number of prior culpable conducts. This objection then observes that in many jurisdictions conviction is a matter for the jury while sentencing is a matter for judge. This objection claims, accordingly, to impose separate convictions for the prior faults entails that the jury has to consider a lot more offences, which is inefficient. What we should do, instead, is to allow the judges to take those prior faults into account in sentencing for the present crime, according to this objection.

In reply, this objection’s force will be blunted once we remember that minor offences are tried by judges instead of the jury in most common law systems. Setting that point aside, once we realise that we can create broad-brushed multiple-conduct offences to cover a large number of individual prior culpable conducts by a few convictions, then even if the prior faults are treated separately from the present crime caused by them in the conviction stage and even if the offences covering the prior faults are considered by the jury, it does not entail that the jury has to consider a lot more offences. In order to illustrate this point, consider an addicted thief whose addiction was caused by his culpably trying drugs on day 1, culpably taking drugs again on day 3, culpably taking drugs again on day 4, etc. If we impose a separate conviction for each instance of drug-taking, then the jury may need to consider a lot more offences. But instead of doing that, we can simply charge him with a broad-brushed offence of “culpable use of drug through a particular period”. In such a case, the jury would only need to decide one more charge and it is unclear that substantial efficiency must be sacrificed.

4. The student exam case revisited

I wish now to revisit the student exam case, which I mentioned in section 1 of this chapter above. Although reprimands are different from criminal punishment, *they are also expressive*. Hence I believe the three expressive considerations in criminal punishment I raised in section 3.3.2 above for my proposal, with suitable modifications, also apply to the student exam case. I will suggest how the student exam case should be properly dealt with by applying those considerations.

Recall the student exam case. Compare two students who did *equally poorly* in an exam. One of them (let us call him the rich student), however, received a very good academic education whereas the other (let us call him the poor student) did not. Let us stipulate that the rich student failed to acquire the mental states that were necessary for doing well in the exam (say, the correct mathematical beliefs) because of his prior *fault* (in a sense of fault relevant for academic institutions, which need not be the same sense of fault relevant for criminal punishment) – say, if he made a normal effort during term time he would have acquired the relevant academic qualities, but that is not the case for the poor student. How should this be reflected in the associated prizes and reprimands?

Some people believe that the poor student should be blamed less *for his poor performance* and this should be reflected in the associated prizes and reprimands. For example, suppose the school gives out reprimands for poor exam performance. It might be suggested that if the rich student gets a first-degree reprimand for his poor performance in the exam, the poor student should get a second-degree (less severe) reprimand for his poor exam performance.

I do *not* believe that is the right way of dealing with the students. I think if the rich student receives a first-degree reprimand *for his poor performance*, so should the poor student. This does not mean that the school should treat the students in the same way *in all aspects*: I believe the rich student should receive *a separate reprimand for his lack of effort* (or the poor student should receive a separate token of appreciation for his effort despite the fact that he did not do well in the exam). This way of dealing with the students is better than simply giving the poor student a less severe reprimand for his poor exam performance for three reasons (which are analogous to the three considerations I offered, in section 3.3.2 above, for why we should not punish an offender, who is indirectly blameworthy for his present crime, more for his present crime but instead we should impose separate convictions and punishments for his prior faults). First, the rich student's lack of effort would be more clearly marked by the separate reprimand concerning effort (or the poor student's effort would be more clearly recognised by the separate token of appreciation). Second, the time of the rich student's fault (or the poor student's good conduct) would be more clearly marked. This information would be important if we want to know about the progress of the students. Third, if we separate the reprimand *for lack of effort* from the reprimand *for poor exam performance*, it would be clear what a first-degree reprimand for exam performance means: it means a very poor exam performance. In contrast, if we take into account students' efforts in giving out reprimands for exam performances, then the messages sent out in the reprimands for exam performance become more ambiguous.

I believe my proposed way of dealing with the student exam case is quite plausible. The academic prizes and reprimands in many schools are entirely based on academic merit. Good effort, or the lack of it, is recognised separately through things like good progress awards or verbal praises. If my proposal is plausible in dealing with the student exam case, then that provides indirect support for my rejection of the prior fault principle in criminal punishment.

5. Conclusion

In this chapter, I have criticised the prior fault argument, which claims that, first, if an offender was not culpable for acquiring the mental states that motivated his crime, then he is less blameworthy for that crime and hence there is a reason to punish him less for that crime, and second, the proportion of socially deprived offenders who were not culpable for acquiring the mental states that motivated their crimes is greater than that for those who are non-socially deprived offenders. I have granted the empirical claim of the argument and I have tried to reject the moral claim of the argument.

Chapter 4: Moral education and direct blameworthiness

1. Introduction

In the last chapter (chapter 3), I have explored *the prior fault argument*, a historical exculpatory argument which appeals to the prior fault principle. According to the prior fault principle, if an offender is not culpable for acquiring the mental states that motivated his crime then he is less blameworthy for his crime, and accordingly, there is a reason to punish him less for the crime than an otherwise similar offender who was culpable for acquiring the mental states that motivated his crime. Under the prior fault principle, a history of bad moral education or upbringing, which socially deprived offenders generally have, *only* has an *instrumental* significance: if an offender had a bad moral education, then it may be *more likely* that he was not (or at least less) culpable for acquiring the mental states that motivated his crime. I have tried to reject the prior fault principle in the last chapter.

Whether or not my assessment of the prior fault principle is satisfactory, some people may believe that a history of bad moral education should have an *intrinsic*, rather than (or in addition to) instrumental, significance to an offender's blameworthiness for his crime, and accordingly, to how much punishment he should receive for the crime. To them,[122] an offender with a bad moral education is less blameworthy for his crime, and accordingly there is a reason to punish him less for his crime, than an otherwise similar offender who had a good moral education, *even if it is not true* that the crime of the offender with a good moral education is traceable to his prior faults; and this, in turn, provides a basis, *independent from the prior fault argument,* for claiming that there is a partial excuse that is more applicable to socially deprived offenders.

I believe deciding whether this argument, which I will call *the education argument*, is sound is quite practically important. In Canada and Australia,[123] the courts held that an offender with a bad moral education and upbringing is less blameworthy for his crime, and accordingly there is a reason to punish him less for his crime than an offender with a good moral education, even if they are similar at the time of offence. But the courts there do *not* inquire as to whether the crimes by offenders with a bad moral education are more, or rather less, traceable to their prior faults as compared to crimes by offenders with a good upbringing. This may suggest that the rationale for

---

[122] The distinction between arguments based on bad moral education and arguments based on lack of prior faults is not often drawn in the literature, but there are a number of writers who argue that wrongdoers with a bad moral education are less blameworthy for their wrongdoings without making much use of the concept of prior fault. See, e.g., Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances"; Campbell, "Is Free Will a Pseudo-problem?", 477; Ashworth, "Justifying the Grounds of Mitigation", 9.

[123] *Bugmy* [2013] HCA 37; *Ipeelee* [2012] 1 S.C.R. 433. See also the US capital sentencing decisions, e.g. *Penry,* 492 U.S. 302 (1989), 319.

punishing offenders with a bad moral education less for their crimes, in those jurisdictions, is *simply because* they had a bad moral education, *regardless of* whether their crimes are less traceable to their prior faults. Assessing, and ultimately *rejecting*, the education argument is the main task of this chapter. Let me begin by laying out the two main claims of the education argument.

The education argument

Moral claim: If an offender had a bad moral education then he is less blameworthy for his crime, and accordingly, there is a reason to punish him less for his crime, than an otherwise similar offender who had a good moral education, even if the crime of the latter offender is not (more) traceable to his prior faults.

Empirical claim: The proportion of socially deprived offenders who had a bad moral education is greater than that for those who are non-socially deprived offenders.

Before assessing the education argument, I wish to note four points, which are important to bear in mind when we assess it.

First, by a good “moral education” or “upbringing”, as I mentioned in chapter 3, I refer to an upbringing that is conducive (i.e. makes it easier and more likely) to one's acquisition of mental states that are conducive to moral and law-abiding behaviour. Having a bad moral education, accordingly, is criminogenic. A good moral education involves both cognitive and conative components: it refers to an environment that is conducive to the acquisition of, in addition to correct moral beliefs, other relevant mental states like proper desires, just like a good "musical upbringing" refers to an environment that encourages the acquisition of both the correct beliefs about music and an appetite for good music.

Second, the education argument, *like the prior fault argument,* is a *historical* argument. According to the education argument, an offender with a bad moral education or upbringing is, *because of his history,* less blameworthy for his crime than an offender with a good moral education *even if the two offenders were exactly identical at the time of offence.*

Third, recall the distinction between *direct* and *indirect* blameworthiness, which I introduced in chapter 3. A person's indirect (or derivative) blameworthiness for a conduct depends on whether and to what extent the conduct is a result of his other prior conducts for which he is directly blameworthy and how directly blameworthy he is for those other prior conducts. Since the education argument claims that an offender who had a good moral education is more blameworthy for his crime (than an otherwise similar offender who had a bad moral education) *even if his crime is not traceable to his prior faults*, the education argument must be based on the idea of *direct* blameworthiness: that an offender with a bad moral education is, because of his history, less *directly* blameworthy for his crime. This point is worth emphasizing. The claim that an agent's *indirect* blameworthiness for a conduct can depend on history is not controversial in the literature on free will and moral responsibility.[124] The education argument, however, is making the much more controversial point that an agent's *direct* blameworthiness for a conduct can depend on history. When I talk about blameworthiness below in this chapter, I am *always referring to direct* blameworthiness rather than indirect blameworthiness unless the context clearly suggests otherwise.

Fourth, I wish to stress that the education argument is an *exculpatory* argument, i.e. an argument that claims that there is a reason to punish some offenders less for their crimes *on the ground that they are less morally blameworthy (i.e. deserves less severe negative moral reactive attitudes) for those crimes*. In this chapter, I will argue that

[124] Bratman, "Fischer and Ravizza on Moral Responsibility and History". Even most so-called non-historicists about moral responsibility accept the idea of tracing, i.e. the idea that indirect blameworthiness depends on history.

the education argument is unsuccessful as it is not true that an offender with a bad moral education *is less blameworthy for his crime* than an otherwise similar offender who had a good moral education. But even if my assessment of the education argument in this chapter is correct, I *would not yet have shown* that there is no reason to *punish* an offender with a bad moral education less for his crime because I will *not*, in this chapter, discuss the issue of whether bad moral education can be relevant to punishment *in a non-exculpatory manner*, i.e. the issue of whether there is a reason to punish an offender with a bad moral education less for his crime *even if he is not less blameworthy for his crime.* It is at least conceptually possible to argue that while an offender who had a bad moral education is not less blameworthy for his crime, given his defective upbringing it is less fair to ask him to bear the consequences of his blameworthy choice. I will consider (and ultimately reject) the non-exculpatory argument that focuses on moral education along this line, which will be called the "education and burdens argument", in chapter 9.

I will grant the empirical claim of the education argument and focus on assessing the moral claim of the education argument in the rest of this chapter. The remaining parts of this chapter will be arranged as follows. In section 2, I will lay out my view on moral blameworthiness, *the quality of will view.*[125] (I do not claim originality for it since it is a popular view in the literature on moral responsibility.) According to this view, a person's moral blameworthiness for a conduct depends on, in addition to the wrongfulness of the conduct, two factors. First, whether he is a member of the moral community, by that I mean whether he can understand moral considerations in general. A person who is not a member of the moral community, like an infant, is exempted from all moral responsibility. Second, his quality of will, i.e. his concern for others, when he acted. If an agent acted with a lesser degree of concern for others when he performed his conduct, then he is more morally blameworthy for it. Since a person's quality of will in performing an act does not (non-instrumentally) depend on his upbringing, if the quality of will view is true, then a person's moral blameworthiness for an act would not depend on his upbringing.

In section 3, which forms the main part of this chapter, I will explore, and ultimately reject, a defence of the moral claim of the education argument that appeals to a different view of moral blameworthiness, which can be called *the variable expectation view* of moral blameworthiness. The main difference between the quality of will view and the variable expectation view is that according to the variable

[125] The classic statement of this view can be found in Strawson, "Freedom and Resentment". Recent discussions can be found in, e.g., McKenna, *Conversation and Responsibility*; Smith, "Culpable Ignorance"; Scanlon, *What We Owe to Each Other,* chapter 6.

expectation view, an agent's moral blameworthiness for an act does *not* simply depend on, *judged by an absolute standard,* how bad is his quality of will in performing the act.[126] Rather, it depends on *how far* his quality of will *falls below* the standard of good will we can expect from him and *our expectations of good will should vary depending on a person's difficulty (capacity and opportunity) to act morally*. If we accept the variable expectation view on moral blameworthiness, and if it is also true that a person who had a bad moral education had less opportunity to act morally even if his circumstances and constitution were identical at the time of offence, then it follows that offenders with a good moral education are more morally blameworthy for their crimes (even if their crimes are not traceable to their prior faults). In response, I will reject the variable expectation view on moral blameworthiness.

2. The quality of will view of moral blameworthiness and two historicist objections

According to the quality of will view of moral blameworthiness, a person's moral blameworthiness for an act depends, besides the wrongfulness of the act, on two

[126] I ignore pleas based on exemptions and reduced wrongfulness here.

factors: first of all, whether he understands moral considerations in general; second, his degree of ill will, i.e. his lack of concern for others' interests and rights, in acting. (In what follows I am referring to cases where there is no difference in wrongfulness.)

Let us start with the moral understanding requirement. By saying that an agent needs to understand moral considerations in order to be blameworthy for a wrongful act, I do not mean that he has to understand *that particular act* to be morally wrong. I mean, rather, that he must have some participant understanding of moral considerations *in general*; when others criticize him for acting immorally, he must have an internal understanding of the kind of attitude or emotion they are expressing in doing so. The idea that if an agent does not understand any moral considerations at all, then he is exempted from moral blame, is shared by many.[127] This moral understanding requirement can be supported both theoretically and by considering some particular cases.

Theoretically, it is plausible to believe that moral blame is, at least incipiently, expressive.[128] In order for an expressive act to make sense, it must be possible for the recipient to understand the blaming message, even if he will not agree with that.

[127] Shoemaker, "Moral Address, Moral Responsibility, and the Boundaries of the Moral Community", 70; Watson, "Responsibility and the Limits of Evil"; McKenna, *Conversation and Responsibility.*
[128] Watson, "Responsibility and the Limits of Evil".

Accordingly, if an agent does not understand moral considerations in general, he cannot be a suitable target of blame. An additional, but closely related, theoretical rationale for the moral understanding requirement is offered by McKenna.[129] According to McKenna, blame in a dimension is appropriate only towards an act that *bears meaning in that dimension*. The act by a person who entirely lacks understanding of a certain dimension cannot bear meaning concerning that dimension. For example, consider an ancient samurai who lacks any understanding of chess at all – he does not even have an idea that it is a game. In such a case, even if he moves the pieces in a certain manner, his act does not bear any *chess meaning*. (His act *may bear meaning in dimensions other than chess* and may invite reactive attitudes in those dimensions. Suppose the reason why he moved the pieces was because he thought this "ritual" could kill his father and he could therefore inherit his father's property. In such a case, his act may bear *moral meaning* and is thus an appropriate target of *moral* blame – my point is only that it would not bear *chess meaning*.) It would be senseless, then, to assess whether the move is a good chess move and target him with reactive attitudes based on that. It might be thought that, analogously, the acts of an agent who is devoid of any understanding of morality cannot bear any moral meaning at all, and hence, it would be senseless to assess his act morally. The moral

[129] McKenna, *Conversation and Responsibility,* 85-104.

understanding requirement can also be supported by considering some particular cases. It explains neatly why toddlers and animals are not blameworthy for their acts, even if they acted with insufficient concern for others' interests and rights.[130]

Much more can be said about the moral understanding requirement and exemptions. However, exemptions are *not* the main focus in this chapter. This is because the focus of this chapter is on the moral blameworthiness of offenders with a bad moral education and offenders with a bad moral education are not necessarily creatures that have no participant understanding of moral considerations in general like animals; so let us move on from *exemptions* to *excuses*.[131] According to the quality of will view, if a person can understand moral considerations in general, then his conducts are open to moral assessment. How morally blameworthy he is for a wrongful conduct, and accordingly whether he has a full or partial moral excuse, depends on his quality of will in performing the conduct.[132] A person's quality of will is worse *if he cares less about the interests and rights of others.* For example, if I would kill a stranger if I am paid $100, but another person would never kill a stranger for money no matter how

130 Strawson, "Freedom and Resentment".
131 For the terms "exemptions" and "excuses", see Strawson, "Freedom and Resentment".
132 This is a very popular and influential account. See, e.g. Strawson, "Freedom and Resentment"; Hieronymi, "The Force and Fairness of Blame"; Scanlon, *What we Owe to Each Other*; Talbert, "Implanted Desires, Self-formation, and Blame"; Mckenna, *Conversation and Moral Responsibility*; Arpaly, *Unprincipled Virtue*; Alexander, "Insufficient Concern: A Unified Conception of Criminal Culpability".

much money he would get by doing so, then other things being equal that person has a better quality of will than me.[133]

The idea that a non-exempted agent's moral blameworthiness for a wrongdoing depends on his quality of will in committing the act seems to be supported, again, both by theoretical reasons and by consideration of particular cases. Moral reactive attitudes, like indignation and resentment, seem to be distinguished from non-moral reactive attitudes precisely because they are concerned with other people's interests and rights. For example, reactive attitudes concerned with prudence or aesthetics seem precisely to be distinguished from moral blame because moral blame is about interests and rights of others.[134] Moreover, the quality of will view can explain many of our considered judgments about particular cases. It can explain why agents who act under factual ignorance should sometimes be excused and sometimes not. An agent who killed under the mistaken factual belief that he was shooting a table (instead of a human) should be excused but a gangster who made a mistaken factual belief *about the identity of the victim only* should not be excused. The quality of will view can explain this: the first agent did not act with lack of care of interests or rights of others, but killing a person, no matter the identity of the victim, reveals a lack of concern for

[133] Arpaly, *Unprincipled Virtue*, 86.
[134] Scanlon, *What We Owe to Each Other,* 287.

interests and rights of others. The quality of will view can also account for when threats exculpate and when threats do not have that effect. A person who robbed a bank *because of a death threat* has an excuse but a similar robber who faced an equal threat but who did not rob the bank *because* of the threat should not be excused. The quality of will view can explain this: the first robber did not act with insufficient care about the interests and rights of others, but the same cannot be said for the latter robber.[135]

If the quality of will view of moral blameworthiness is correct, then the moral claim of the education argument must be false. It is because one's quality of will in acting depends only on his *present* features, i.e. the extent to which he was motivated by interests and rights of others *in acting*. This is a separate issue from the historical issue of whether he had a good moral education and upbringing.[136]

According to the quality of will view, direct blameworthiness is a non-historical matter. In the next section, I will explore a competing account of direct blameworthiness, which is (or at least can be) [137] historical, namely, the variable

---

[135] This is a similar case as the overdetermined killer case, which I discussed in chapter 2 of this thesis above.

[136] Ebels-Duggan, "Dealing with the Past"; Hieronymi, "The Force and Fairness of Blame"; Talbert, "Implanted Desires, Self-formation, and Blame".

[137] Conceptually speaking, the variable expectation view can be historical or non-historical, depending on whether one takes capacity and opportunity to involve historical components. (This is different from

expectation view. But before doing so, I want to explore two historicist objections that are frequently raised against any non-historical accounts of direct blameworthiness, including the quality of will view.

The first objection, which can be called *the manipulation objection*, appeals to our intuition concerning cases where an agent's motivating mental states were directly instilled by some third party, like a case where a mad scientist injected a drug into the agent's brain causing him to acquire a completely new set of mental states (values, desires, beliefs, etc.) for a few hours, which in turn motivated his wrongdoing.[138] According to the proponents of the manipulation objection, the agents in the manipulation cases are not blameworthy for their acts but a non-manipulated agent who acted under the same mental states is. However, this verdict is, according to this objection, inconsistent with any non-historical account of direct blameworthiness.

I want to make two replies to the manipulation objection. The first reply is that, even for someone who accepts a non-historical account of direct blameworthiness, there are two justifications that he can appeal to in explaining why in *some* cases the non-manipulated agent is more blameworthy for his act. First, some non-manipulated

the quality of will view, which must be non-historical.)

[138] E.g. Fischer and Ravizza, *Responsibility and Control*, chapters 7 and 8.

agents are *indirectly* blameworthy for their acts. Given that a manipulated agent is clearly not indirectly blameworthy for his act (barring exceptional cases where, for example, the agent had invited the manipulation), we can sometimes explain the difference in the manipulated agent's blameworthiness for his act and the non-manipulated agent's blameworthiness for his act by appealing to a difference in their indirect blameworthiness for their acts without accepting that direct blameworthiness depends on history.[139] Second, it may be argued that the reason why some manipulated agents are not blameworthy for their acts is *not because of their histories preceding their acts*. Rather, it is because the agents whom we are now trying to hold responsible are so psychologically disconnected with the agents who performed the acts.[140] To illustrate, according to the second justification, the reason why the manipulated agent in the injection case mentioned above is not blameworthy for his act is not because of his history preceding the act but because of what happened after the act: he, *after the effect of the drug expired*, is radically disconnected from the agent who acted under the influence of the drug. (I will revisit the idea of psychological connectedness in section 3.2 of this chapter below.)[141] These two justifications, I acknowledge, will *not* cover *all* possible manipulation

[139] McKenna, "Responsibility and Globally Manipulated Agents".
[140] Khoury , "Synchronic and Diachronic Responsibility". For a similar reply, see Arpaly, *Unprincipled Virtue*, 163-168.
[141] Manipulation cases offered in the literature frequently share this feature of reversal. For example, in the case discussed by Robinson in arguing for the manipulation defence, the agent changed back to his normal self after committing his crime: Robinson, "Are we Responsible for Who we Are?", 67.

cases. Some non-manipulated agents are not indirectly blameworthy for their wrongdoings (i.e. their wrongdoings are not traceable to other faults of them in the past). The first justification cannot justify the conclusion that a manipulated agent is less blameworthy for his wrongdoing than these non-manipulated agents. Some manipulated agents also did not change back to their previous selves after the manipulation – imagine that the effect of the injection is permanent. The second justification cannot justify the conclusion that such manipulated agents are not blameworthy for their acts. But I am not sure if that is a problem. In cases where the two justifications do not apply, I no longer have a strong intuition that the manipulated agent is less blameworthy for his act. To illustrate, think about someone who was changed into an evil person *permanently* because of an injection by a mad scientist. Then he did something wrong. As Scanlon and many others have argued, I cannot see why this offender deserves less severe negative reactive attitudes than an otherwise similar offender who acquired his evil mental states in an innocent but a less radical (i.e. non-manipulated) manner.[142]

My second reply to the manipulation objection is that even if we accept the soundness of the manipulation objection, the help it can offer to the education argument is

[142] Scanlon, *What We Owe to Each Other,* 279; Frankfurt, "Reply to Fischer"; McKenna, "A Hard-line Reply to Pereboom's Four-case Manipulation Argument".

limited. As King correctly observes, the intuitive cases that the manipulation objection appeals to typically involve some special features: they are cases involving a third party manipulator, and the change in the agent's mind set was sudden.[143] But the cases that are covered by the moral claim of the education argument are cases where these special features are generally absent: even for offenders who had a bad moral education, their mental states generally were not instilled into them through a radical process by a third party.[144] In other words, *even if* the manipulation objection can successfully show that direct blameworthiness is historical, it does not show that direct blameworthiness generally depends on one's moral education, *a particular kind of history,* and hence, the manipulation objection does not offer much support for the moral claim of the education argument even if it is sound. This explains why I think a proponent of the education argument should focus, not on the manipulation objection, but rather on the variable expectation view of blameworthiness, since the variable expectation view can be concerned with *the kind of history* that is relevant for his purpose.

The second historicist objection against the quality of will account is the standard incompatibilist objection: in order for an agent to be blameworthy for an act, it is

143 King, "The Structure of Responsibility", 153-171.
144 For similar points, see Kaye, "The Secret Politics of the Compatibilist Criminal Law", 385-389.

insufficient that he acted with a bad quality of will; in addition, his bad quality of will must not be causally determined by the past. I have two replies. First, as many compatibilists have argued, it is hard to see why causal determination is inconsistent with blameworthiness.[145] Second, similar to my second reply to the manipulation objection, the help that the incompatibilist objection can offer to the education argument is limited, even if the objection itself is sound. It is because, as Moore has argued, it is hard to see on what basis we can say that the bad quality of will of offenders with a bad moral education is causally determined in a way that the bad quality of will of offenders with a good moral education is not.[146] If the world is causally determined, then the quality of will of everyone is causally determined by the past. Accordingly, even if the incompatibilist objection can show that some kind of history matters, it would not show that a history of bad moral education matters. That explains why I do not think it is necessary to launch a detailed investigation as to whether causal determination is consistent with moral responsibility in my thesis.

3. The variable expectation view of moral blameworthiness

[145] See, e.g., Wolf, *Freedom within Reason,* 103-106.

[146] Buss, "Justified Wrongdoing", 339; Moore, "Causation and the Excuses"; Morse, "Deprivation and Desert". For a reply that I do not regard to be successful, see Kaye, "Resurrecting the Causal Theory of the Excuses"; Kaye, "The Secret Politics of the Compatibilist Criminal Law", 403-405.

In this section, I wish to explore (and ultimately reject) an alternative view of moral blameworthiness, *the variable expectation view*, which may be invoked to support the moral claim of the education argument. The variable expectation view is not a view radically different from the quality of will view. (I take this to be a virtue of the variable expectation view since I think the quality of will view is quite attractive so even if it is not ultimately correct it is implausible to believe that the ultimately correct view of moral blameworthiness would be radically different from it.) The variable expectation view does not disagree with the quality of will view insofar as exemptions are concerned. The variable expectation view also agrees with the quality of will view that, *other things being equal*, if a non-exempted agent acted with a worse quality of will in committing a wrongful act then he is more morally blameworthy for the act.

The main difference between the two views is that according to the variable expectation view, an agent's moral blameworthiness for a wrongdoing *does not simply* depend on the quality of his will in acting, *judged against an absolute standard;* rather, it depends on *how far his quality of will fell below our normative expectation*, and our expectation of good will *should be lowered if it is more difficult for the agent* *to act morally.*[147] Following Nelkin, *I define difficulty to act morally as depending on one's capacity and opportunity to act morally*: it is more difficult for a person to X if he has less capacity or opportunity to X. I understand *capacity* as referring to reason-responsiveness and reactivity.[148] How much opportunity a person had to X depends on whether (and on how many occasions) he could X by making certain choices, how much cost and effort were needed for him to make those choices, and his knowledge of the consequences of his decisions, etc.[149] (Though I wish to observe that I do not think my discussion in this chapter depends on any particularly controversial understanding of the concepts of capacity and opportunity.)

Many people believe that the concept of opportunity has both present and *historical* components: a chess player who had years of good training had more opportunity to play well in a game than a novice player *even if his situation and constitution were identical to the novice player at the time of playing that game*, because he had more opportunity to acquire the skills in the past which would enable him to play well in the game. If the variable expectation view is true, and if it is also true that the concept

[147] For versions of this view, see Nelkin, "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness"; Kelly, "Culpability and Fault in Criminal Justice"; Franklin, "A Theory of the Normative Force of Pleas"; Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances"; Smith, "Control, Responsibility, and Moral Assessment", 390; Smith, "Responsibility for Attitudes", 268-269.

[148] Fischer and Ravizza, *Responsibility and Control*; Duff, *Answering for Crime*, 39.

[149] For similar accounts of opportunity, see Nelkin, "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness"; Arneson, "Equality and Equal Opportunity for Welfare", 56; Scanlon, *Moral Dimensions*, 202-205.

of opportunity has a historical component such that an offender who had a bad moral education, because of his history, had less opportunity to act morally, then we should lower our expectation of good will from him.[150] Accordingly, even if he acted with the same quality of will at the time of offence, he is less blameworthy for his act since his level of good will *fell less short* of the standard of good will we expect from him.[151]

It should be noted that the variable expectation view, *per se*, does *not* entail the moral claim of the education argument. Accordingly, one can reject my opponent's reasoning in two ways: first, by rejecting the variable expectation view on moral blameworthiness; second, by accepting the variable expectation view on moral blameworthiness but rejecting the claim that an offender with a bad moral education had less capacity or opportunity to act morally (for example, one may hold a non-historical account of capacity and opportunity such that how much capacity and opportunity one has for acting morally only depends on his situation at the time of performing the act). I will only pursue the *first* route in attacking my opponent's reasoning, i.e. I will grant that an offender with a bad moral education, *because of his*

[150] Here I assume that capacity is by and large a non-historical concept.
[151] Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances". For similar thoughts, see Vogel, "Understanding and Blaming", 139; Wolf, *Freedom within Reason*, 86-87, 121-122.

history, had less opportunity to act morally than an otherwise similar offender with a good education and focus on arguing that the variable expectation view of moral blameworthiness is incorrect.

I wish to stress one point about the significance of the concept of difficulty under the quality of will view before we move on. Even under the quality of will view, difficulty may seem to matter to moral blameworthiness in certain cases. The quality of will view can agree that it is more difficult for an agent who was under a death threat to act morally, and if he did the wrong act because of the threat, then he is less morally blameworthy for the act. But under the quality of will view, the importance of difficulty is *entirely evidential or instrumental*: the reason why difficulty matters *in some cases* is *because it bears on the agent's quality of will in those cases*. According to the quality of will view, if the death threat did not affect the agent's quality of will in committing the act (say in over-determined incentive cases) then the extra difficulty would not matter to the agent's moral blameworthiness for his act.[152] In contrast, the variable expectation view is giving the concept of difficulty an *intrinsic* role: the variable expectation view claims that for two agents who acted *with the same quality of will,* one of them is less morally blameworthy for his act if it is more difficult for

[152] Talbert, "Moral Competence, Moral Blame, and Protest", 102.

him to act morally. Whether an agent's difficulty to act morally can affect his direct moral blameworthiness *without affecting his quality of will* is the issue dividing the variable expectation view and the quality of will view.

There is certainly some intuitive appeal to the variable expectation view – we sometimes say we should tailor standards and expectations for short-comers.[153] But when we examine the view closer, its normative basis becomes more suspect. The variable expectation view is not calling for discarding all the expectations of good will, as what we would do when we face animals or toddlers. What it claims is that we should *lower* our expectation of good will for some people depending on their difficulty to act morally. The implication is that instead of caring about how good or bad a person's quality of will is, judged absolutely, we should care about *the difference* between the agent's quality of will and a certain variable standard set by his opportunity and capacity. But while I can see the moral significance of one's actual degree of ill will, I cannot see the moral significance of the *difference in amount* between the actual degree of ill will and some variable standard. It obviously matters to me whether you are a person with a good will; it does not, at least not obviously, matter to me how your degree of good will fare against a certain benchmark set by

[153] See, e.g. Horder, *Excusing Crime*, 108, chapter 4.

your individual capacity and opportunity.

It may be tempting to defend the idea that we should adjust normative expectations according to difficulty by appealing to the “ought implies can” principle.[154] But that would seem to beg the question here because a defender of the quality of will view, like me, is precisely rejecting the “ought implies can” principle in the relevant sense: the quality of will view holds that a person’s moral blameworthiness for doing something only depends on his quality of will in doing that, and insofar as whether a person “can” do otherwise does not bear on his quality of will in acting, it is irrelevant to his blameworthiness for that conduct.[155]

I do not have a direct knockdown argument against the variable expectation view but I wish to launch an indirect attack against it. Many people who are drawn to the variable expectation view do so because they think the variable expectation view can accommodate for our intuition concerning certain cases. In fact, many writers who hold the variable expectation view explicitly do so based on consideration of certain cases.[156] If the quality of will view can account equally well or even better for our

[154] Ciurria, “Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances”, 553.
[155] See Frankfurt, “Alternate Possibilities and Moral Responsibility”; Talbert, “Moral Competence, Moral Blame, and Protest”.
[156] E.g. Kelly, “Culpability and Fault in Criminal Justice”; Franklin, “A Theory of the Normative Force

intuition in those cases, then that would constitute a power case against the variable expectation view. That is what I intend to do in sections 3.1 to sections 3.3. The three cases I wish to consider concern youngsters who act under provocation, wrongdoers who act under emotional stress, and novice chess players who made bad moves in chess.

3.1. Provocation and youth

Consider two people who both attacked their victims under provocation but one of the provoked attackers was a youngster while the other one was an adult. Many people believe that the youngster is less blameworthy for losing his self-control if the gravity of provocation is similar. This case may offer some support to the variable expectation view as the variable expectation view seems to be able to account for the case. Youngsters have less capacity to maintain their self-control, and hence less capacity to act morally, in the face of provocation. Therefore, we can explain why youngsters are less morally blameworthy for their conducts in provocation cases by claiming that we lower our expectation of good will for them in virtue of their lowered capacity to act

---

of Pleas"; Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances".

morally under provocation.[157]

In reply, I agree that youngsters, as compared to adults, generally have less capacity to act morally under provocation. I also accept that many (but *not all*, as we will see below) youngsters are less morally blameworthy for their wrongdoings under provocation if the gravity of provocation is similar. But I do *not* believe that their diminished capacity is the ground for excuse in those cases, and accordingly, I do not believe that a consideration of provocation cases supports the idea that we adjust our expectation of good will depending on one's capacity to act morally. Rather, I will argue that the reason why many youngsters are less morally blameworthy for their acts under provocation is *simply because their quality of will in acting is less bad, judged by an absolute standard*. I will defend my suggestion in two parts. First, I will explain why many youngsters who act wrongly under provocation act with a less bad quality of will. Second, I will argue that a youngster who had less capacity to act morally but who acted with an equally bad quality of will should not be exculpated.

There are two reasons why many provoked youngsters commit their wrongdoings with a less bad quality of will than provoked adults. First, youngsters may not be as

[157] For discussion of related cases, see Horder, *Excusing Crime*, 109.

consciously aware of the factual consequences of their acts in attacking.[158] Part of the process of growing up is to acquire the ability to represent more fully the consequences of one's act in one's conscious deliberation. As such while a provoked youngster who attacked may have full conscious awareness of the *immediate* consequences of his action (say, the victim's leg will bleed) it is less likely for him to be as consciously aware of the *further* consequences of his act (say, the long term impact on the victim and the victim's family, etc.) as a provoked adult. If moral blameworthiness depends on the concern one has for others in acting, and one's concern for others in acting depends partly on what one consciously takes to be the bad consequences of one's act,[159] then many provoked young attackers may act with less disregard for interests or rights of others as compared to provoked adult attackers.

Second, even assuming that a provoked youngster had the same conscious awareness of the factual consequences of his act when he attacked, it is still possible for him to have acted with a less bad quality of will. It might be objected that a person's quality of will in acting depends only on what desire he acts for and what he takes to be the factual consequences of satisfying the desire; since the provoked youngster tried to satisfy the same desire with his act as the provoked adult (e.g. "to give the victim

[158] Scott and Steinberg, "Blaming Youth".
[159] Husak, "Negligence, Belief, Blame and Criminal Liability", 215-216.

what he deserves"), if, as I am assuming here, they had the same foresight of the factual consequences of their acts, then their quality of will in acting must be the same. This objection, however, assumes too narrow a view on what matters to one's quality of will.

A person can avoid acting immorally in two ways. First, one may simply lack desires that would lead him to act immorally. But that of course is not always possible. The second way in which a person can act morally is by *resisting* desires that would lead to immoral actions. A person who cares more about the interests and rights of others would make more effort to resist such desires. As such, *even if* two agents acted immorally for the same desire (e.g. "to give the victim what he deserves") with equally full awareness of the harm they were causing, *their quality of will in acting may differ* depending on *how much (failed) effort* they made to resist their immoral desires in acting.[160]

While effort is necessary for resisting a desire, whether one can successfully resist a desire depends *not only* on one's degree of effort made. It also depends, as the psychological literature on ego depletion demonstrates, on one's amount of

[160] Garvey, "Dealing with Wayward Desire"; Kant, *Lectures on Ethics,* 63.

psychological resources.[161] For example, a person who has not slept for days has fewer psychological resources for self-control. Even if he made the same effort as us in trying to control his desires, he is less likely to succeed. In this aspect, self-control has frequently been said to be similar to muscles: whether one can successfully move one's body depends, in addition to one's effort exerted, on one's muscle strength.[162] A body builder may be able to stand on his arm for ten minutes but I will not be able to do that *even if I tried as hard as the body builder*.

Youngsters generally have fewer psychological self-control resources.[163] That may explain why in many cases we are less inclined to accuse a provoked youngster for not making an effort to control himself, and accordingly for acting with as bad a quality of will, as compared to a provoked adult. To see this point, consider a particular taunt. The gravity of this taunt is one such that an adult, given his psychological resources, will be able to maintain his self-control in the face of it with a reasonable amount of effort. Given that youngsters have fewer self-control resources, a youngster can only maintain self-control in the face of such taunt with *more than* a reasonable amount of effort – *a reasonable amount of effort would not be enough*

[161] Holton, "How is Strength of Will Possible?"; Holton and Shute, "Self-control in the Modern Provocation Defence".
[162] Holton, "How is Strength of Will Possible?"
[163] Scott and Steinberg, "Blaming Youth", 816; Scott, "Adolescent and the Regulation of Youth Crime", 343-345.

given his deficiency of psychological resources. Now suppose both a youngster and an adult were provoked to attack because of the taunt. The adult, in such a case, *clearly* did not make a reasonable effort to control himself. However, we cannot immediately say the same for the youngster – it is *possible* for him to have made a reasonable effort, yet the effort failed because of his deficiency of psychological resources.

Based on the two points above (foresight of factual consequences and effort to control oneself), the quality of will view can explain why many youngsters are less morally blameworthy for their attacks under provocation than adults. Now I want to, based on a modification of a thought experiment originally offered by Frankfurt, argue that youth, and accordingly reduced capacity to act morally, does *not* affect an agent's moral blameworthiness for his act under provocation *if it does not bear on his quality of will* in acting.[164]

Consider an adult and a youngster who both faced the same taunt as the one I mentioned two paragraphs above. The adult did not make a reasonable effort to resist his rage and as a result he was provoked to lose his self-control and attacked. The youngster, given his lack of psychological resources, would not be able to resist his

[164] See Frankfurt's discussion of "willing addicts" in *The Importance of What We Care About*, 24-25, 51-52, 96; for similar ideas, see Tadros, *Criminal Responsibility*, 62.

rage even if he has made a reasonable effort; yet, in the case we are considering the youngster did not even *bother to make any effort to resist his rage* at all – he simply *gave in* to the rage. Moreover, he had full vivid awareness of the consequences of his conduct throughout.

It seems to me that the youngster in this case is *not* less morally blameworthy for attacking than the adult. But the variable expectation view would predict otherwise. It is because, under the variable expectation view, how morally blameworthy a person is depends on whether and how far his level of good will fell short of the level we can expect from him, and the level of good will we can expect from a person *depends on his capacity to act morally*. Given that the youngster had less capacity to control himself and to act morally, he must be less morally blameworthy for his act if his quality of will in acting is the same. The quality of will view, in contrast, would predict the right conclusion in this case: since the youngster did not make any extra effort to resist the immoral desires, he did not care more about the interests and rights of others, and accordingly, he is not less morally blameworthy for attacking than the provoked adult.[165]

[165] This approach is broadly in line with Tadros's discussion in *Criminal Responsibility,* 36-37.

3.2. Emotional stress

The next case I wish to consider, which I take from Franklin, concerns agents who act under emotional stress.[166] Think about Smith, who normally is a nice person, but who recently is under emotional stress because of his job search. One day, in a restaurant, he made a really rude remark to his friend. Many people believe that Smith is less morally blameworthy for making the rude remark than someone who is not under emotional stress. This case may offer some support to the variable expectation view as the variable expectation view seems to be able to account for the case. It is plausible to believe that it is more difficult for Smith to act morally because of his stress. Therefore, we can explain why Smith is less morally blameworthy for making the remark by claiming that we set a lower expectation of good will for him because of his extra difficulty to act morally.

I agree that it is more difficult for agents who are under emotional stress to act morally. I also accept that many (but not all, as we will see below) agents who act under emotional stress are less morally blameworthy for their acts. But I will argue that the real basis for why they are less morally blameworthy for their acts is not

[166] Franklin, “A Theory of the Normative Force of Pleas”.

because it is more difficult for them to act morally but because of something else. I will first argue that we can appeal to two considerations, namely, quality of will and psychological connectedness, to explain why some agents who act under emotional stress are less morally blameworthy for their acts without appealing to the concept of difficulty. I will then argue that an agent who is under emotional stress (and accordingly has more difficulty to act morally) but who acts with the same quality of will and whose act is not psychologically disconnected from himself is not less morally blameworthy for his act than someone who does not act under emotional stress; hence considering emotionally stressed agents does not provide any real support for the variable expectation view.[167]

I believe that in many cases emotionally stressed agents may act with a less bad quality of will. The reasons are quite similar to those I raised concerning provoked youngsters in section 3.1 so I will be brief here to avoid repetition. First, agents who act under emotional stress may have less vivid conscious awareness of the factual consequences of their acts. To illustrate using the Smith case above, if a person is under stress about job-seeking then job-seeking would occupy a centre stage in his

[167] I believe my treatment of this case is quite similar to the approach taken in Arpaly, *Unprincipled Virtue*, 149-159.

phenomenal awareness.[168] Smith may not consciously know that his words would hurt the feeling of the listener to the same extent because his mind is largely occupied by job seeking. Second, an agent who is under emotional stress has fewer psychological resources to resist immoral desires.[169] So the fact that he did not successfully resist the immoral desires does not entail that he did not make a reasonable effort to resist them, and accordingly, does not entail that he lacked sufficient concern for others.

The second consideration for why agents who act under emotional stress are sometimes less morally blameworthy for their acts requires a minor *supplement* to the quality of will view. (I discussed a related point in section 2 of this chapter above in reply to the manipulation objection.) Recall that according to the quality of will view, a non-exempted agent's moral blameworthiness for his act *simply* depends on his quality of will in performing the act. But many people think that a person's moral blameworthiness for an act depends, in addition to his quality of will in performing the act, the *psychological connectedness* between the act *and the agent whom we are now blaming*. That explains why many people feel uncomfortable to blame a totally reformed person for a wrongdoing that he committed thirty years ago – the agent no

[168] See Watson, "Disordered Appetites", 71-73, for similar points in relation to addiction.
[169] Holton and Shute, "Self-control in the Modern Provocation Defence", 56.

longer seems very connected to the act. That may provide another basis for why many agents who act under emotional stress are less morally blameworthy for their acts – since their behaviours under stress are usually very different from their normal behaviours.[170] Of course, if we add the psychological connectedness component to moral blameworthiness, then strictly speaking we are modifying the quality of will view instead of accepting it as it is. But it is important to notice that adding the psychological connectedness component does not affect the standard of good will we expect from a person when he acted; accordingly, it provides no support to the variable expectation view.

By appealing to the ideas of quality of will and psychological connectedness, we can explain why many agents who act under emotional stress are less morally blameworthy for their acts without accepting the variable expectation view. Now I want to argue that an agent is not less morally blameworthy for his act simply because he was under emotional stress when he acted (and accordingly had a greater difficulty to act morally) if the stress bears neither on his quality of will in acting nor his psychological connectedness with his act. To see this, consider another person, Sean. Sean, unlike Smith, has always been a rude person who frequently makes nasty

[170] Maybe that is what Strawson means when he writes that an agent who acted under stress should be excused because "he wasn't himself". Strawson, "Freedom and Resentment", 51.

remarks. Now it is near the end of the job-hunting season and he felt stressed. One day, he felt a desire to make a rude remark; he decided not to resist the desire at all and just made the remark. When he made the remark, he knew it would hurt the feeling of others; in fact, that was his intention in making the remark.

It seems intuitive that Sean, unlike Smith, is equally morally blameworthy for making the rude remark as a person who made the same rude remark with the same motive without being under emotional stress. This provides an objection against the variable expectation view because that view would predict otherwise. This is because, according to the variable expectation view, our expectation of good will for a person is lowered if it is more difficult for him to act morally; since Sean was under emotional stress and a person who is under emotional stress has a greater difficulty to act morally, he must be less morally blameworthy for making the rude remark if his quality of will when making the remark was the same, according to the variable expectation view.

3.3. Chess and training

The next case we are going to consider is one *not* concerned with *moral* blameworthiness; yet, according to the proponents of the variable expectation view of moral blameworthiness, considering this case provides powerful *indirect* support for their view on moral blameworthiness.[171]

Consider a chess player who made a terrible move. Many people believe that whether and the degree to which he is blameworthy for the bad move depends on his chess training. If the player is a *novice*, he is not, or at least less, blameworthy for the bad move. In contrast, if the player is *an experienced player coached by the best teachers,* then he is (more) blameworthy for the bad move. It is plausible to believe that a novice chess player had less opportunity to play well because he had less opportunity to acquire the skills in the past which would enable him to play well now. It may be thought that *the reason why* novice chess players are less blameworthy for their bad moves is because we judge them by a lower standard of skill, in virtue of their reduced opportunity to play well. If that is the case, then, analogously, it seems plausible that we should lower our expectation of good will for people who have reduced opportunity to act morally, in virtue of their lack of opportunity to internalise moral norms in the past.

[171] Ciurria, "Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances", 52-53. The actual example she used is one about tennis but the difference between chess and tennis should not matter for her point. I chose chess to illustrate simply because I am more familiar with chess so I can tailor my own example, which I will give after discussing hers, better.

I agree that a novice chess player had less opportunity to acquire chess skills, and accordingly to play well, than an experienced player who has received good training. I also accept that many (but not all, as we will see) novice players are less blameworthy for their bad moves. But I do *not* believe the reason why many novice players are less blameworthy for their bad moves is *because* we adjust normative expectation of skill depending on opportunity, and accordingly, on education and training. There are, instead, two *other* explanations that explain why many novice players are less blameworthy for their bad moves.

First, experienced players are often more skilful. If a skilful player made an obviously terrible move in a game it is likely to be due to his lack of *concern* or care when he was playing the game, and lack of *concern* in playing is of course a kind of fault relevant to chess in addition to lack of *skill*. But even if we focus on experienced players who *lacked skill when they made the bad moves* such that their bad moves were not due to their lack of care at the time of playing, we can explain why some of them are more blameworthy for their moves than novice players without accepting that blameworthiness depends on opportunity. We can simply explain that by appealing to their *indirect* blameworthiness: an experienced player may be *indirectly* blameworthy for his bad play in a way that the novice player is not. A person's indirect blameworthiness for a conduct, to recall, depends on to what extent the conduct is traceable to other prior conducts for which he is directly blameworthy and how directly blameworthy he is for those prior conducts. If a chess player had ten years of good training but still failed to grasp some elementary chess principles then his bad play, unlike the same play by the novice, may be traceable to his laziness, or other faults relevant to chess, within his past ten training years.

I wish to stress that the second explanation for why many experienced players are more blameworthy for their bad moves, which appeals to their *indirect* blameworthiness for their bad moves, albeit also historical, is *different from* claiming that a player's blameworthiness for a bad move depends on how much opportunity to play well did he have. This is because the issue of whether and the degree to which *one is indirectly blameworthy for not playing well* is distinct from the issue of *how much opportunity he had to play well*. Two persons can have equal opportunity to avoid X without being equally indirectly blameworthy for X. Suppose two persons voluntarily took drugs, and then both attacked someone under a state of automatism a few hours later. The first person took the drug *recklessly* without intending to do harm. The second person took the drug *with the purpose of increasing his risk of causing* *harm* afterwards. In such a case, the second agent is more indirectly blameworthy for the attack (since his motive in taking the drug is worse), but they may have equal opportunity to avoid the attack (since the cost and effort needed to make the choice not to take the drug may be the same for them and they had same knowledge, etc.). More importantly, two persons can also be equally indirectly blameworthy for a bad outcome while having an unequal amount of opportunity for avoiding the bad outcome. This is because the failure to make use of an opportunity to avoid a bad outcome in the future need not be blameworthy. (This point applies no matter what kind of bad outcome we are talking about and no matter whether we are talking about moral blameworthiness or other kinds of blameworthiness). To see this point, imagine the following scenario. Two adults maliciously injured their victims using their martial arts skills. They would not have committed their crimes if they did not know martial arts. They are identical except the following. The first person was forced to learn martial arts by his parents when he was young. The second person learnt martial arts voluntarily when he was young. The second person knew, when he was young, that by learning martial arts he increased his risk of assaulting someone in the future. But the increase in risk is minimal; he also had a socially acceptable reason to take the risk (say, because of benefits to his health). In such a case, the second person had more opportunity to avoid committing his present crime (since he could avoid the crime by not learning martial arts when he was young, unlike the first person); but he is *not* more indirectly blameworthy for his present crime since his prior conduct, of learning martial arts, *is not a blameworthy one.*[172]

My approach, which explains why *some* experienced players are more blameworthy for their bad moves by appealing to the two ideas of indirect blameworthiness and care about making the good moves, rather than the idea that normative expectation of skill varies depending on opportunity to play well, *will not show that all* experienced players are more blameworthy for their bad moves. But I take that to be a strength, rather than a weakness, of my approach as education and opportunity to play well *per se* do not seem to affect a player's blameworthiness for a bad move.

To see this, consider an experienced player with a good chess education who made a bad endgame move in a tournament game. He cared about winning and tried his best to think of the best move when he was playing, but he could not find the correct move at that time. Let us also stipulate that he had an opportunity to avoid the bad move in the past: had he spent more training time on his endgame he would be able to come up with the correct move on the tournament day. However, the amount of training we can

[172] For further illustrations of the idea that culpability and opportunity are entirely distinct concepts, see Scanlon, *Moral Dimensions*, 201; Tadros, *The Ends of Harm*.

reasonably expect from a player is finite and while it is possible for this player to work harder during his training, he was not at fault in not doing so.

Compare this experienced player with a novice player who made an equally bad endgame move. Is the experienced player more blameworthy for the bad move than the novice? I find it hard to see how we can target him with more negative reactive attitudes for the move. *What can we fault him for, besides the lack of skill at the time of playing (which is a fault also displayed by the novice player)?* My proposal predicts the correct verdict: this experienced player deserves equal direct blame in terms of his skill in playing, equal direct blame in terms of his concern in playing, and equal indirect blame, so he is equally blameworthy for the move as the novice player. However, according to my opponent’s proposal, this experienced player is more blameworthy for his bad move. This is because, according to my opponent’s proposal, our normative expectation of endgame skill varies depending on the player’s opportunity to play skilfully; given that the experienced player played with an equally bad endgame skill on the tournament day, but he had more opportunity to develop the endgame skills and accordingly to play a skilful endgame, he is more blameworthy for the unskilful move.

Since the chess and training case is concerned with the issue of whether a bad chess education affects chess blameworthiness, an issue closely connected to the main issue of this chapter, which is whether a bad moral education affects moral blameworthiness, it may be worthwhile to explicitly recap my view concerning both cases side by side. Let us start with chess. I believe that some experienced players with a good chess education are more *indirectly* blameworthy for their bad chess moves as their bad plays are traceable to their faults in training years. Moreover, some experienced players may deserve more blame for their bad moves if they cared less *at the time of playing*. But the importance of chess education to chess blameworthiness is *entirely evidential* – if, firstly, the bad move by the experienced player is not traceable to his faults in the past and secondly, his state of mind (and accordingly his degree of skill and care) was the same at the time of making the move as the novice, then the experienced player is not more blameworthy for the bad chess move simply because he had a better chess education.

My position concerning the relationship between moral blameworthiness and moral education is analogous. I believe that some offenders with a good moral education are more indirectly blameworthy for their present crimes, i.e. their present crimes are more traceable to their other culpable conducts in the past. Moreover, some offenders with a good moral education may act with a worse quality of will when they committed their present crimes. But the importance of moral education to moral blameworthiness, just like the importance of chess education to chess blameworthiness, is *entirely evidential*: if, firstly, the present crime by an offender who had a good moral education is not (more) traceable to his prior faults and secondly, he was identical to the offender with a bad moral education at the time of committing his present crime (and accordingly acted with the same quality of will), then he is not more morally blameworthy for his present crime simply because he had a better moral education; accordingly, we should not accept the moral claim of the education argument.

4. Conclusion

In this chapter, I have examined the education argument, which is a historical (i.e. an argument that does not appeal to the features of the offenders at the time of offence) exculpatory argument.

According to this argument, socially deprived offenders are more likely to have a bad moral education; and if an offender had a bad moral education then he is less morally blameworthy for his crime, and accordingly there is a reason to punish him less for his crime, than an otherwise similar offender who had a good moral education, even if the crime of the latter is not (more) traceable to his prior faults. I focused on examining, and ultimately rejecting, the moral claim of the education argument.

I first provided a brief defence of the quality of will view of moral blameworthiness, a view of moral blameworthiness that is inconsistent with the moral claim of the education argument. I then examined another view of moral blameworthiness, which can be called the variable expectation view and which can be used to defend the moral claim of the education argument. I have argued that, despite the intuitive appeal of the variable expectation view of moral blameworthiness, we should reject it.

Chapter 5: Non-culpable normative mistake

1. Introduction

Sarah Buss argued that socially deprived offenders are often less morally blameworthy for their crimes, and accordingly, there is often a reason to punish a socially deprived offender less for his crime, because socially deprived offenders often act under a non-culpable mistaken belief in committing their crimes.[173] Assessing this line of thought, which appeals to the concept of non-culpable mistake, is the task of this chapter.

Let me start by sharpening this line of thought. There are two types of mistaken belief: normative mistakes and factual mistakes. Factual mistakes are mistakes about the physical state of affairs of the world, e.g. I thought a person was carrying a gun while in fact he was carrying a toy. Normative mistakes are one's mistakes about what he ought to do given his beliefs about the factual circumstances.[174] For example, a

[173] Buss, "Justified Wrongdoing". This kind of argument is also discussed in Morse, "Deprivation and Desert", 149-150. See also Wolf, "Sanity and the Metaphysics of Responsibility", 61.

[174] Normative mistakes can be further divided into two kinds: first, mistakes about the requirements of morality, and second, mistakes about the reason-giving force of morality. I will not distinguish between these two kinds of normative mistake in this chapter. For the distinction, see, e.g., Buss, "Justified Wrongdoing", 358; Rosen, "Culpability and Ignorance", 75. Milo nicely discussed the distinction and its moral significance in "Wickedness".

murderer who shot his victim to death may be under no mistake about the factual circumstances of his action: he knew that he was pressing the trigger of a gun, that the person in front of him would be killed by his action, etc. His only mistake was normative, as to what action he had most reasons to do given his factual beliefs.

Most writers who argue that offenders with a certain background are often less blameworthy for their crimes, by appealing to the idea of non-culpable mistake, focus on normative mistakes rather than factual mistakes.[175] While they are seldom explicit about their rationales for this focus, the rationales are not hard to see. First, it is unlikely that a substantial portion of serious wrongdoings are committed under factual mistakes. Second, the claim that socially deprived people are more likely to acquire mistaken *normative* beliefs non-culpably has some plausibility. In contrast, it seems less plausible to hold that non-socially deprived agents who act under *factual* mistakes are more likely to have acquired their mistaken beliefs culpably. Imagine two policemen who both shot their victims under the mistaken factual belief that the victim was attacking them with firearms. One of the policemen is socially deprived and the other is not. I do not think there is any obvious argument for claiming that the policeman who was not socially deprived was more likely to have acquired the

175 E.g. Morse, "Severe Environmental Deprivation", 160, 163; Morse, "Deprivation and Desert", 149-150; Levy, "Cultural Membership and Moral Responsibility".

mistaken factual belief culpably.

In what follows, I will *set aside agents who acted under factual mistakes* and focus on examining whether there is an argument for claiming that there is a partial excuse that is more applicable to socially deprived offenders by appealing to the concept of non-culpable *normative* mistake. Let me lay out the argument that I will assess (and ultimately reject) in this chapter, which can be called the *non-culpable normative mistake argument*, before going on:

The Non-culpable Normative Mistake Argument

Moral claim: If an offender committed his crime under a mistaken normative belief that was acquired non-culpably, then he is less blameworthy for his crime and accordingly there is a reason to punish him less for his crime than an otherwise similar offender who did not commit his crime under a non-culpable normative mistake.

Empirical claim: The proportion of socially deprived offenders who committed their crimes under a mistaken normative belief that was acquired non-culpably is greater than that for those who are non-socially deprived offenders.

I will grant the empirical claim of the non-culpable normative mistake argument for the sake of discussion and focus on the moral claim of the argument. How can an offender commit a crime without acting under a non-culpable normative mistake? As mentioned above, I will set aside offenders who act under factual mistakes. Offenders who do not commit their crimes under a non-culpable normative mistake (and who do not commit their crimes under a factual mistake) fall under *two* categories. First, offenders who act under a *culpable* normative mistake. Second, offenders who act under no mistaken belief at all: these offenders know that, all things considered, they ought not commit their wrongdoings; but nonetheless they act against their all things considered judgment about what they should do. These offenders act under what most writers refer to as akrasia (or weakness of the will, which I will take as interchangeable).[176] In order for the moral claim of the non-culpable normative mistake argument to be true, an offender who committed his crime under a non-culpable normative mistake must be less blameworthy for his crime, and

[176] I use both terms, "weakness of the will" and "akrasia", to refer to cases where an agent acts voluntarily against his belief that he should do otherwise all things considered. I understand that some writers draw a distinction between the two terms and they use "weakness of the will" to refer, not to actions against one's belief as to what one should do all things considered, but to actions that involve an "over-ready" revision of one's previous intention. See Holton, "Intention and Weakness of Will"; Rosen, "Skepticism about Moral Responsibility", 309. Since Holton's conception of weakness of the will is consistent with acting under a mistaken normative belief (I may revise my intention in an over-ready fashion because of a belief), I am focusing on akrasia rather than weakness of the will in Holton's terminology.

As a side note, it is very unclear how a wrongdoer who acted under weakness of the will in Holton's sense must be more blameworthy than a wrongdoer who did not.

accordingly there is a reason to punish him less for his crime, than *both* categories of offenders.

Concerning the *first category* of offenders who did not act under a non-culpable normative mistake (i.e. offenders who acted under a *culpable* normative mistake), I have already argued in chapter 3 (the chapter on the prior fault argument) that there is no reason to punish an offender who acquired the mental states that motivated his crime culpably more *for that crime* than an otherwise similar offender who acquired the mental states that motivated his crime non-culpably. Since mistaken normative beliefs are mental states, it *follows* from my argument in chapter 3 that there is no reason to punish an offender who committed his crime under a culpable normative mistake more *for that crime* than an otherwise similar offender who committed his crime under a non-culpable normative mistake. To recap my main point briefly, I have argued that while there is a reason to punish the offender who committed his crime under a normative mistake that was acquired culpably more *in total*, there is no reason to punish him more *for that crime*; instead, there is only a reason to *impose separate convictions and punishments for his prior faults* that led to the normative mistake motivating his present crime.

In this chapter, I want to raise a different challenge to the moral claim of the non-culpable normative mistake argument by focusing on the *second* category of offenders who did not commit their crimes under a non-culpable normative mistake (i.e. offenders who acted under weakness of the will). I will argue against the idea that an offender who committed his crime under a non-culpable normative mistake is less blameworthy for his crime than an otherwise similar offender who acted under weakness of the will.[177] In fact, the position on the moral status of akratic wrongdoers that the non-culpable normative mistake argument is committed to (and which I am arguing against) can be called the *revisionary position* as it goes against the *traditional position* on the topic: writers since Aristotle usually hold that vice is worse than akrasia, even if the vice was non-culpably acquired.[178] I will first lay out an argument for the traditional position on akratic wrongdoers in section 2. I will then survey and reject arguments for the revisionary position in section 3.

2. The traditional position on akratic wrongdoers

As mentioned, the non-culpable normative mistake argument, which compares wrongdoers who act under a non-culpable normative mistake favourably to akratic

[177] For people who accept such an idea, see Rosen, "Culpability and Ignorance". For a very similar point, see Zimmerman, "Moral Responsibility and Ignorance".
[178] Aristotle, *Nicomachean Ethics*, book VII, 1151a25.

wrongdoers, goes against the traditional position on akratic wrongdoers, which holds that akratic wrongdoers are actually less blameworthy for their crimes than mistaken wrongdoers. The traditional position is quite intuitive and one can defend it by appealing to the *wholeheartedness argument* offered by Arpaly and Schroeder. According to this argument, a person's blameworthiness for a wrongdoing depends on the extent to which the psychological features motivating the wrongdoing were integrated within his overall personality. A wrongdoer who acted under a normative mistake need have no conflict in her psychology; in contrast, the desire that motivated an akratic wrongdoer must be in conflict with another part of his psychology, namely, his normative belief.[179] Isn't it intuitive to believe that a wholehearted devotion to wrongfulness is, other things being equal, worse than a half-hearted devotion?

3. The arguments for the revisionary position on akratic wrongdoers

I will consider (and ultimately reject) two arguments, which can be called the *factual mistake analogy argument* and the *reasonable expectation argument* respectively, for the revisionary position. The *factual mistake analogy argument* starts by making the uncontroversial point that non-culpable *factual* mistakes generally exculpate. By

[179] Arpaly and Schroeder, "Praise, Blame, and the Whole Self".

analogy, since non-culpable factual mistakes exculpate, non-culpable normative mistakes must exculpate as well. Since a wrongdoer who acts under a non-culpable normative mistake is not blameworthy, but a wrongdoer who acts akratically is, the revisionary position is established.

The problem of this argument is well-known: there is a perfectly natural justification for why factual mistakes exculpate, which does not apply to normative mistakes. The justification is based on the *quality of will* of the agent: an agent who acted under a factual mistake may act with sufficient concern for interests of others, while acting under a normative mistake, like the belief that the interests of other people do not count or the belief that women are mere sexual objects, is *constitutive* of the fact that the agent acted with insufficient concern for the interests of others even if the mistaken belief had been non-culpably acquired. So without further arguments, it is unmotivated to claim that if factual mistakes exculpate then so must normative mistakes.[180]

In fact, if we focus just on the *quality of will* of the agent in committing his act, it is hard to see why the non-akratic offender is less blameworthy for his act. To illustrate,

[180] Rosen, "Culpability and Ignorance", 72-73; Fields, "Moral Beliefs and Blameworthiness".

a rapist who acted with the belief that women are inferior and hence raping them is acceptable does not seem to act with a lesser lack of concern for the interest of the victim than another rapist who knew he should not rape all things considered but succumbed to his sexual desires.[181] If anything, as the wholeheartedness objection suggests, the non-akratic rapist would seem to act with a greater lack of concern for his victim, since in addition to being motivated by immoral desires, he is also motivated by immoral beliefs.

The second argument for the revisionary position, the *reasonable expectation argument*, which is offered by Rosen, starts with the idea that in order for an agent to be morally blameworthy for an act, it is *insufficient* for him to act with a bad quality of will. *In addition*, it must be *reasonable to expect* him to act otherwise. This argument continues by claiming that *it is unreasonable to expect a non-culpably mistaken agent to act otherwise*, accordingly, an offender who committed his wrongdoing under a non-culpable normative mistake is not blameworthy for his conduct, even if he acted with a bad quality of will. Since an akratic offender is blameworthy for his conduct but an offender who acted under a non-culpable

[181] Talbert, "Moral Competence, Moral Blame, and Protest", 101; "Unwitting Wrongdoers and the Role of Disagreement in Blame"; Tadros, *Criminal Responsibility,* 324-326.

normative mistake is not, the revisionary position is established.[182]

I have argued in the last chapter that a person's (direct) moral blameworthiness for a wrongdoing depends, in addition to the wrongfulness of the act and the psychological connectedness between the act and the agent whom we are trying to hold responsible, on two considerations: first, whether he can understand moral considerations in general; second, his quality of will in committing his act. If that is correct, then reasonable expectation, *insofar as* it is a concept distinct from those two considerations, is irrelevant to moral blameworthiness. But I am willing to grant here, for the sake of argument only, that my view on blameworthiness defended in the last chapter is mistaken. Let me grant, for the sake of argument, that there is *some* sense of reasonable expectation such that (1) in some cases it is unreasonable to expect a person who acts with a bad quality of will to act morally even though he can understand moral considerations in general; (2) in such a case he is not blameworthy for his act because it is unreasonable to expect him to act otherwise. I believe Rosen's argument fails, even if we proceed on that concession on my part, because he cannot find *a plausible morally significant sense* of reasonable expectation under which it is unreasonable to expect an agent who had a non-culpable normative mistake to act

[182] Rosen, "Culpability and Ignorance", 74-75.

morally.

Wallace believes that if a person lacked general moral competence or capacity, then it is unreasonable to expect him to act morally.[183] But as Rosen himself acknowledges, this point by Wallace would not assist his argument much.[184] This is because general capacity persists over time and one retains the general capacity to do well in a practice even when, at a particular moment, he thought he should not do well in that practice or he made a mistake as to what are the demands of that practice. A professional pianist retained the general capacity to play well at a particular time despite the fact that he thought he should not play well or he happened to make a mistake as to which pedal he should press at that time. Similarly, a normatively mistaken wrongdoer can certainly retain the general capacity to act morally when he acted even though he made a mistake as to whether he should obey the demands of morality or what the demands of morality are.

It is sometimes suggested that it is unreasonable to expect a person to comply with a demand even if he has the general capacity to do so if it is too *costly* for him to do so: it is unreasonable to expect a person under a death threat to act morally. (I do not

[183] Wallace, *Responsibility and the Moral Sentiments*.
[184] Rosen, "Culpability and Ignorance", 74.

think cost matters to morally blameworthiness if it does not affect the agent's quality of will,[185] but I will set this point aside.) But even if we grant that cost is relevant to reasonable expectation and accordingly to moral blameworthiness, it cannot help my opponent here: it need not be costly at all for an agent who made a normative mistake to avoid wrongdoing. It need not be very costly for the normatively mistaken rapist to avoid raping – it may just be that, given that he thought women have an inferior status, he thought it is not worthwhile to pay the cost (by refraining from raping) however small the cost it is.

If an offender who made a non-culpable normative mistake could possess general moral capacity and it need not be too costly for him to do otherwise, then why is it unreasonable to expect him to act morally? Levy appeals to the concept of *rationality.* He claims that it is unreasonable to expect a person to do something that amounts to an irrational conduct, and whether it is rational for one to do X depends on whether he can do X "by means of a reasoning procedure operating over his existing beliefs".[186]

I do not think the concept of rationality helps to establish the conclusion that it is unreasonable to expect agents who made a non-culpable normative mistake to act

[185] Think about the overdetermined cases, for example.
[186] Levy, *Hard Luck,* 128.

morally, and my reply to Levy, which is largely taken from Robichaud, starts by drawing a distinction between two things we can mean by saying that a person can do X rationally, i.e. "by means of a reasoning procedure operating over his existing beliefs".[187]

According to the first interpretation, to say that a person can do X by means of a reasoning procedure operating over his existing beliefs means that *he can reach the conclusion to do X via a sound deliberative route based on his existing beliefs*.[188] This interpretation of rationality cannot help Levy as it may be rational for a normatively mistaken agent to act morally under this interpretation. While a normatively mistaken agent lacks the present belief that he should act morally, it is entirely possible for there to be resources in his existing beliefs that can lead him to the conclusion that he should act morally via a sound deliberative route.[189] As an analogy, I now believe that death penalty is unacceptable; but it is possible for there to be other beliefs in my belief system, which I do not notice at present, that can lead me

[187] Robichaud, "On Culpable Ignorance and Akrasia".
[188] Williams defines the concept of "internal reasons" using a similar idea. See Williams, "Internal Reasons and the Obscurity of Blame", 35.
[189] Robichaud, "On Culpable Ignorance and Akrasia". The objection I raised against Levy in the main text is that it may be rational for a mistaken agent to act otherwise under the first sense of rationality, so Levy's argument does not work even if we assume that in order for an agent to be blameworthy for an act, it must be rational for him to act otherwise under the first sense of rationality. For the record, let me also observe that I do not believe that in order for an agent to be blameworthy for an act, it must be rational for him to act otherwise under the first sense of rationality, i.e. I do not believe that in order for an agent to be blameworthy for an act he must have internal reasons (in Williams's sense) to act otherwise. See Talbert, "Unwitting Wrongdoers and the Role of Disagreement in Blame", 241; Shafer-Landau, *Moral Realism*, 181-183.

to the contrary conclusion via a sound deliberative route.

I also wish to observe that to say that there are resources in the mistaken wrongdoer's belief system that can lead him to the conclusion to do otherwise via a sound deliberative route does *not* mean that his present belief that he should commit the wrongdoing is culpably acquired. There are at least two reasons why a person who believes not-X may reach the conclusion X via a sound deliberative route without being blameworthy for acquiring the belief not-X. First, it is not a requirement of rationality that one should scrutinize all one's reasons before believing something. We often have to form beliefs and act under time constraints.[190] Second, not all failings of rationality are morally blameworthy. If a person is terribly bad in logical reasoning, his rationality is defective. But it is hard to see how that *per se* can attract moral blame.[191] Accordingly, under the first interpretation of rationality, it is possible for it to be rational for a normatively mistaken agent to act morally, even if the mistaken belief is non-culpably acquired.

According to the second interpretation of rationality, a person can do X by means of a reasoning procedure operating over his existing beliefs *only if he, at the time of action,*

---

[190] Montmarquet, "Epistemic Virtue and Doxastic Responsibility", 332; "'Pure' versus 'Practical' Epistemic Justification".
[191] Tadros, *Criminal Responsibility,* 253.

believes he should do X. But if Levy's argument is based on this conception of reasoning procedure and rationality, then he seems to just beg the question. It is unclear why a person who is not otherwise persuaded by Levy will accept that it is unreasonable to expect a person to do X unless doing X is rational in this *particular* sense of rationality.[192]

Besides "rationality", "agency" is a concept that Levy appeals to in defending the claim that we cannot reasonably expect normatively mistaken agents to act morally. Levy claims that it is unreasonable to expect a person to do X if X amounts to "a glitch of agency".[193] The claim that we cannot expect someone do to an act that amounts to a glitch of agency in some sense seems plausible: I can forget something only as a result of a glitch of agency, and we do not usually blame people for failing to forget something. Akrasia is also usually thought to be a glitch in agency. Does it not follow that we cannot reasonably expect a normatively mistaken agent to act morally, since to do so is to expect him to act akratically, which is an act that amounts to a glitch of agency?

The problem of this argument is that it trades on an ambiguity of the phrase "glitch of

[192] Robichaud, "On Culpable Ignorance and Akrasia".
[193] Levy, *Hard Luck,* 128.

agency". We can distinguish between two understandings of the phrase. Sometimes we use the phrase to refer to some unwelcomed behaviours *outside one's control*. For example, suppose I got hiccups when I gave a public speech. That is a glitch in my agency – I cannot control whether I had hiccups or not. Forgetting something also amounts to a glitch in agency in this sense. According to the second understanding of the phrase, I have a glitch in agency if I do not act *optimally* in accordance with my conscious reasons at that time. I agree that it is unreasonable to expect someone to perform an act that amounts to a glitch in agency in the first sense – it is unreasonable to expect someone to do something that he can't control. However, akratic actions do not involve glitch in agency in the first sense – they only involve glitch in agency in the second sense as whether one performs an akratic action is something within his control (that it is within our control whether to perform akratic actions is precisely what distinguishes akratic actions from *compulsive* actions).

The claim that a wrongdoer who acted under a non-culpable normative mistake cannot be reasonably expected to act morally may also draw support from the following line of reasoning. It starts by claiming that if a person acquired his mistaken beliefs non-culpably, then it is unreasonable to expect him to form different beliefs. The reasoning continues by claiming that *if it is unreasonable to expect someone to do*

X, then it would also be unreasonable to expect him to do Y, if Y is a result of X. Now consider the rapist who acted under a mistaken normative belief that was acquired non-culpably. It is unreasonable to expect him to form a different belief. How then is it reasonable to expect him to act morally, given that his immoral action *was a result of his belief*?

In reply, there are two problems with the above line of reasoning, which relies crucially on the principle that "if it is unreasonable to expect someone to do X, then it would also be unreasonable to expect him to do Y, if Y is a result of X". First, if this principle is valid then it would be unreasonable to expect many akratic wrongdoers to act morally as well. Think about akratic wrongdoers who were *non-culpable* in acquiring their weak wills. Since it is unreasonable to expect them to acquire a stronger will, it follows from the principle that it would also be unreasonable to expect them to act morally since their akratic wrongdoings are a result of their weak wills. In other words, the principle cannot draw a distinction between normatively mistaken wrongdoers who acquired their mistaken beliefs non-culpably and akratic wrongdoers who acquired their weak wills *non-culpably*, and thus, fails to establish the revisionary conclusion that a normatively mistaken wrongdoer is less blameworthy for his act than an akratic wrongdoer. Second, the principle is actually just a variation of a famous principle frequently offered in support of incompatibilism in the free will literature, the principle that "if a person is not blameworthy for X and if Y is a result of X then he cannot be blameworthy for Y" in disguise.[194] The problem of the principle is well known. The fact that my choice is a result of the past, which I have no control over (and accordingly which I cannot be blameworthy for), does not entail, without further justifications, that I cannot be blameworthy for my choice. Analogously, I cannot see why the fact that we cannot reasonably expect a person to modify his beliefs entails that we cannot reasonably expect him to make a different choice even if his choice was determined by his beliefs.

Based on the above, I believe that the reasonable expectation argument does not work since it cannot be defended by appealing to the concepts of general capacity, cost, rationality, or agency, or the principle that "if it is unreasonable to expect someone to do X, then it would also be unreasonable to expect him to do Y, if Y is a result of X". Since neither the factual mistake analogy argument nor the reasonable expectation argument for the revisionary position works and the traditional position is both intuitive and supported by the wholeheartedness argument, I conclude that we should reject the revisionary position.

[194] For discussion of this principle, see Fischer and Ravizza, *Responsibility and Control*, 152.

4. Conclusion

In this chapter, I have examined the non-culpable normative mistake argument. According to the argument, the proportion of socially deprived offenders who commit their crimes under a non-culpable normative mistake is greater than that for those who are non-socially deprived offenders; and if an offender committed his crime under a non-culpable normative mistake, then he is less blameworthy for his crime, and accordingly there is a reason to punish him less for the crime, than an otherwise similar offender who did not commit his crime under a non-culpable normative mistake. I have tried to reject the moral claim of this argument.

It might be helpful to use the following table to recap my position on how a person's punishment for a crime should be affected by different kinds of mistake:

Is there a reason to punish an offender less *for the crime* done under a mistaken belief than an otherwise similar offender who did not act under any mistake, i.e. an akratic offender?	The mistake is factual	The mistake is normative
The agent was not culpable for acquiring the mistake	**Yes**	**No**
The agent was culpable for acquiring the mistake	**Yes**	**No**

The reason why the two columns have different answers, i.e. why factual mistakes are treated differently from normative mistakes, is provided in section 3 of this chapter. The reason why the two rows have the same answer, i.e. why a culpable mistake is treated in the same way as a non-culpable mistake *insofar as the punishment for the crime done under the mistake is concerned,* is provided *not* in this chapter but in chapter 3 (on the prior fault argument).

Part II: Partial Justifications

Chapter 6: Natural duties and duty to obey the law

1. Introduction

In this chapter, I will consider two arguments for claiming that there is a partial justification, i.e. a mitigating factor based on reduced wrongfulness, that is more applicable to socially deprived offenders.

The wrongfulness of a crime depends on the stringency and the amount of obligations (or "duties", which I will use interchangeably with "obligations") the offender violates in committing his crime. As an example, compare two men who both touched a third person. The third person consented to the touching by the first man but not the second. The motive of the two men in touching the third person may be the same. But only the second man acted wrongfully in touching as we do not have an obligation to refrain from touching a person who has consented. As another example, compare two men who both destroyed a car. The first man, however, owned the car that was destroyed by him but not so for the second man. In such a case, only the second man's act is wrongful as we do not have a general obligation to refrain from destroying our own property.

It is plausible to believe that if an offender violated fewer or weaker obligations in committing his crime then other things being equal, his crime is less wrongful and accordingly there is a reason to punish him less for the crime. For example, most people believe that a person who touched his victim without consent can be punished but that is not the case if the victim consented. Similarly, most people believe that a man who destroyed his own car should not be punished.

In order to understand the difference between the two arguments discussed in this chapter, let us draw a distinction between natural duties, i.e. duties to do something independent of the command of the law, and the general duty to obey the law.[195] We have some duties that exist independent of legal commands: for example, murder is wrong regardless of whether murder is criminalised. But many people believe that the law can also make a moral difference, that is, many people believe that the law can, at least in certain cases, by its command, create new moral duties. If the law has not

[195] For the distinction between the two kinds of duty, see, e.g., Duff, *Punishment, Communication, and Community*, 56-59, 64-66; Green, "Legal Obligation and Authority"; Simmons, *Moral Principles and Political Obligations*.

specified whether we should drive on the right or drive on the left then we have not committed any wrongdoing by driving on the left. However, if the law has made driving on the left an offence, then many people believe that it would be morally (not simply legally) wrong to do so.[196]

According to the first argument to be considered in this chapter, which can be called the *no violation of natural duties argument*, there is a partial justification that is more applicable to socially deprived offenders because the crimes by socially deprived offenders are more likely to involve no violation of *natural duties*. According to the second argument to be considered in this chapter, which can be called the *general duty to obey the law argument*, there is a partial justification that is more applicable to socially deprived offenders because they are more likely to have a reduced *general content-independent moral duty to obey the law*. I will argue that the no violation of natural duties argument succeeds; however, it can only cover a particular type of offences (namely, property offences against the rich). I will then argue that the general duty to obey the law argument is unsuccessful and that even if it is successful, it is unlikely to justify more than a very modest amount of sentence reduction.

196 Raz, *The Morality of Freedom*, 30.

2. The no violation of natural duties argument

Let me lay out the two claims of the no violation of natural duties argument before discussing it:

The No Violation of Natural Duties Argument

Moral claim: If a person violated no natural duties in committing his crime, then his crime is less wrongful, and accordingly there is a reason to punish him less for his crime than an otherwise similar offender whose crime involves a violation of natural duties.

Empirical claim: The proportion of socially deprived offenders who violated no natural duties in committing their crimes is greater than that for those who are non-socially deprived offenders.

I think the moral claim is unlikely to be controversial and is supported by the consent to touching case and the damage to property case mentioned above.

Let us move on to the empirical claim. I think it is plausible because socially deprived offenders, unlike non-socially deprived offenders, *have no natural duty to refrain from committing a significant category of offences*. The category of offences I have in mind is *purely property offences* (e.g. theft without causing any personal injury) *against rich victims*. (As an example, think about car theft against a rich victim.) If, as it seems plausible, generally speaking the crimes by non-socially deprived offenders involve violation of natural duties, but a significant category of crimes by socially deprived offenders does not, then the proportion of socially deprived offenders who violated no natural duties in committing their crimes is greater than that for those who are non-socially deprived offenders.

The idea that a socially deprived person may violate no natural duty in stealing from a rich victim, as opposed to a theft committed by a non-socially deprived thief, is intuitive. This intuitive idea has also been defended persuasively by Reiman and Green recently.[197] According to them, if a person has benefited because of past wrongdoings that caused loss to another person, then the victim has a right to get redress from the beneficiary. Since the poverty of many socially deprived persons was a result of historical injustice (say, of colonialism or racial discrimination) and the rich people benefitted from it, a socially deprived person has a right to get redress

[197] Green, "Hard Times, Hard Time", 66-69; Reimen, "The Moral Ambivalence of Crime in an Unjust Society". See also Shelby, "Justice, Deviance and the Dark Ghetto", 152.

from the rich people. Morally speaking, therefore, a socially deprived offender is extracting the just compensation that is due to him in stealing a car from a rich person, which he has the natural right to do.[198] In contrast, a non-socially deprived thief has no right to extract compensation from the rich people and accordingly he violated a natural duty in stealing the car.

Let me consider three rejoinders against this argument, which are not considered by Green and Reimen themselves. First, it may be argued that most rich victims of property crimes have not *themselves* committed wrongdoings that led to the disadvantage of the socially deprived thief (the wrongdoings may have been committed by the past generations, or by other more powerful people in the society, such as the capitalist politicians); accordingly the socially deprived thief has no right to obtain redress from them.

In reply, I do not think it is plausible to believe that the victim of a wrongdoing has a right to obtain redress from a beneficiary only if the beneficiary himself committed the wrongdoing. If a robber robbed a victim and deposited the money in my bank

[198] Green, "Hard Times, Hard Time", 66-69; Reimen, "The Moral Ambivalence of Crime in an Unjust Society".

without my knowledge, the victim has the right to ask for his money back.[199]

Second, it may be argued that, insofar as societal injustice is concerned, there are *multiple beneficiaries* and *multiple victims.* While the poor people as a group may have a right to obtain redress from the rich people as a group, there is no reason why this particular victim, namely the socially deprived thief, should obtain redress from this particular beneficiary, namely the victim of the car theft.

In reply, I believe that this rejoinder would *only* have purchase if the rich people have collectively provided *alternative* mechanisms for compensating the poor people, say, by collectively donating into a compensation fund, which the poor people can then sort out how to distribute amongst themselves, or creating a matching scheme regulating which particular poor person(s) should a particular rich person compensate. In absence of such mechanisms, a poor person has a right to extract compensation from any rich person (at least if the amount extracted is not more than the amount that particular rich person needs to contribute in total to discharge his compensatory duty – a condition that is clearly satisfied for most property crimes against rich people).[200] To illustrate, suppose that Ann and Bill owed Carol and David $20. Ann and Bill have,

[199] Both the first rejoinder and my reply are largely taken from Butt, "On Benefitting from Injustice".
[200] For a similar condition, see Murphy, *Moral Demands in Nonideal Theory,* chapter 5.

however, refused to answer the questions by Carol and David as to how or when the debt would be settled for an unreasonably long period of time. Suppose Carol now took $5 from Ann. Given that Ann's group has consistently denied Carol's group just compensation, it seems perverse for Ann to claim that Carol should try to pursue the group instead of her or that she only has a duty to compensate Carol's group but not Carol.

Third, it may be argued that even if a socially deprived thief does not violate any natural duties in taking the car from the rich victim, a non-socially deprived thief *also* would not violate any natural duties in doing so. This is because, given that the existing pattern of possessions is a result of distributive injustice, the rich victim does not have a natural right over "his" car. So no one has any natural duty to refrain from taking the car from him. Accordingly, the argument cannot draw a distinction between socially deprived thieves and non-socially deprived thieves.

In reply, before stating my two reasons for rejecting my opponent's rejoinder, let me first observe that my opponent's rejoinder is counter-intuitive. As an analogy, while a victim of robbery has the right to get his property back from the robber, we do not think other people have a right to take the property from the robber even if the robber is not entitled to it. My two reasons for rejecting the rejoinder are as follows. First, we may think of the rich victim as holding his property on trust for people whom he has a duty to redress (i.e. the socially deprived people).[201] So while he has no natural right concerning the car *against the socially deprived people*, he has a better natural right over the car than the non-socially deprived thief. Second, we may think that there is a natural duty, based on stability, to respect a person's possession of a piece of property unless you are morally entitled to it.[202] That explains why the non-socially deprived thief would violate a natural duty in taking the car from the rich victim, even if the rich victim does not have a natural moral title over the car.

We should bear in mind that the duty not to steal may be justified, not only by natural duties, but also by the duty to obey the law. So I wish to observe that the argument above does not entail (though it is also consistent with the idea) that a theft by a socially deprived offender against a rich victim is not morally wrong *all things considered* – even if the socially deprived offender has a natural right to extract compensation, it may be argued that such a right is weaker than the general duty to obey the law and accordingly he should refrain from stealing, all things considered.[203]

---

[201] For the idea that the rich people hold the property on trust for others, see, e.g. Horne, "Review of *A Discourse on Property*".

[202] For the idea that there is a natural duty to respect possessions based on stability, see Waldron's discussion of Hume's theory of property in Waldron, "'To Bestow Stability upon Possession': Hume's Alternative to Locke".

[203] Reimen, "The Moral Ambivalence of Crime in an Unjust Society", 9-13.

But the argument aims to at least show that property crimes by socially deprived offenders against rich victims violate *fewe*r duties as compared to property crimes by non-socially deprived offenders because the former do not involve violation of *natural* duties.

Let me end by making three points about the scope of the no violation of natural duties argument. First, this argument applies only to property crimes but not other crimes. A person's natural duty to refrain from rape or murder, unlike one's natural duty to refrain from stealing, depends neither on the social condition of himself nor the social condition of the victim.[204] Second, while the argument only applies to property crimes, it is broader than the defence of necessity.[205] Stealing a car generally would not fall under the defence of necessity; but as I have argued, a car theft by a poor offender sometimes would not involve violation of any natural duties. Third, this argument does not, as Green observes, apply to property crimes against *poor victims*. It is because a poor victim has no duty to compensate the socially deprived thief, unlike the rich people, so the socially deprived thief has no right to take the property from a poor victim.[206]

[204] See, e.g. Tadros, "Poverty and Criminal Punishment", 392.
[205] So the argument here is different from Aquinas's argument for claiming that theft can be justified in circumstances of necessity. See Aquinas, *Summa Theologica*, question 66. See also Green, *Lectures on the Principles of Political Obligation*, 146-147.
[206] For discussion of this issue, see Green, "Hard Times, Hard Time", 69.

3. The general duty to obey the law argument

As mentioned above, the no violation of natural duties argument can only cover property crimes but not other types of crime; and as such, that argument may not be entirely satisfactory to those who believe that social deprivation should have a more general effect on wrongfulness of crimes. The general duty to obey the law argument, in contrast, is meant to have a broader application: if it works, it will cover non-property crimes like murder. According to the argument, when a person commits a crime like murder he simultaneously violates two obligations. First, he violates the natural duty to refrain from killing. Second, *in addition*, he violates *the general duty to obey the law*. (The second duty is different from the first duty since it is content-independent.) Since socially deprived offenders are more likely to have a weakened general duty to obey the law, their crimes are often less wrongful even if they have the same natural duty to refrain from performing the relevant acts. Let me lay out the two claims of this argument before assessing it:

The General Duty to Obey the Law Argument

Moral claim: If a person had a reduced general duty to obey the law, then his crime is less wrongful and accordingly there is a reason to punish him less for his crime than an otherwise similar offender who had an uncompromised general duty to obey the law.

Empirical claim: The proportion of socially deprived offenders who had a reduced general duty to obey the law is greater than that for those who are non-socially deprived offenders.

I think the moral claim of the general duty to obey the law argument is quite plausible. As an analogy, think about the duty to keep promises, which is also a content-independent duty like the duty to obey the law. Suppose I promised you not to beat you. Then I beat you. Beating you would, of course, be wrongful even if I did not make the prior promise; but in such a case my act, which involves a simultaneous violation of the duty not to beat people and a content-independent duty to keep promises, seems more wrongful than just beating you without promising you not to beat you. In any case, I will grant for the sake of discussion the moral claim of the general duty to obey the law argument.

Is the empirical claim of the general duty to obey the law argument, which holds that the proportion of socially deprived offenders who have a reduced duty to obey the law is greater than that for those who are non-socially deprived offenders, plausible?[207] In order to answer this, let us first ask: Regarding the *non*-socially deprived offenders, what are the supposed bases to justify the existence of a general duty to obey the law?

The most popular candidates are the following four: the consent theory (the theory that they have a general duty to obey the law because they have consented to that), instrumental justifications like the normal justification thesis (the theory that they have a general duty to obey the law because by doing so they can provide instrumental benefits to others, like security, which they are morally required to provide), receipt of benefit theory (the theory that they have a general duty to obey the law because they have received some important benefits from the state) and the voluntary acceptance of benefit theory (the theory that they have a general duty to obey the law because they voluntarily accepted some important benefits from the state).[208] Whether a person has consented to obey the law and whether his obedience would bring instrumental benefits to others do not seem to depend on whether he is socially deprived; so if the non-socially deprived people's duty to obey the law is

[207] For writers who accept the empirical claim of the general duty to obey the law argument, see, e.g., Hudson, "Beyond Proportionate Punishment", 66-68; Wright, *Does the Law Morally Bind the Poor?*, chapter 1.

[208] For a survey of these theories, see Green, "Legal Obligation and Authority"; Simmons, *Moral Principles and Political Obligations*.

justified on consent or instrumental grounds only, then it is hard to see why socially deprived people are more likely to have a reduced duty to obey the law.[209] Accordingly, the claim that socially deprived offenders are likely to have a reduced general duty to obey the law seems plausible *only if* the receipt of benefit theory or the voluntary acceptance of benefit theory is a basis for justifying a duty to obey the law on the *non*-socially deprived offenders.

Can we justify a general duty to obey the law on the non-socially deprived offenders on the ground of acceptance or receipt of benefits? Before we answer this question, let us observe the difference between the two theories. You can *receive* some benefit without having a reasonable opportunity to avoid it; however, in such a case, you did not *accept* the benefit since acceptance implies reasonable avoidability.[210] Now let us examine the two theories separately starting with the receipt theory.

The receipt of benefit theory faces two serious problems. First, as Nozick argues it is hard to see how one can acquire duties to reciprocate simply by receiving benefits if

[209] Wright has argued against the claim that consent can ground a general duty to obey the law on the socially deprived offenders. See Wright, *Does the Law Morally Bind the Poor,* chapter 1. But even if that is granted, it is unclear how that can show that non-socially deprived offenders are likely to have a stronger general duty to obey the law since it is implausible to believe that many non-socially deprived offenders have provided the relevant consent as well. For objections to the consent justification of the duty to obey the law in general, see Simmons, *Moral Principles and Political Obligations*.

[210] Green, "Legal Obligation and Authority". I take reasonable avoidability to be a necessary condition for acceptance. It may not be sufficient. Simmons seems to think that, in addition, the recipient must have certain psychological attitudes. Whether he is correct does not matter here. Simmons, *Moral Principles and Political Obligations,* chapter 5.

he does not have a chance to avoid them. This is supported by Nozick's famous radio station example: the mere fact that other villagers, without my consent, took turns to run a radio programme broadcasting to everyone in the village, which benefited me, does not create any obligation on me to reciprocate.[211] Second, even if we grant that *some* obligations to reciprocate can arise if an agent received benefits from another without voluntarily accepting them, it is unclear why the obligation created is the duty to obey the benefactor instead of doing other actions to benefit him.[212]

It is plausible to believe that voluntary acceptance of benefit, unlike mere receipt of benefit, would generate some obligations on the recipient. Accordingly, the acceptance of benefit theory avoids the first objection against the receipt of benefit theory. But it faces other problems. Let us begin by asking what exactly is the benefit obtained by the non-socially deprived citizens from the state. Two answers are possible. First, material benefits like wealth. Second, "order and security", etc.[213] If we focus on the first type of goods, then since these goods are reasonably avoidable (if you do not want them you can just refrain from participating in economic activities) it is plausible to say that the non-socially deprived citizens have accepted them rather

[211] Nozick, *Anarchy, State and Utopia,* 93.
[212] Simmons, "The Obligations of Citizens". See also Green, "Legal Obligation and Authority": "Loyalty to one's friends is not normally shown in obeying them. Why should loyalty to the community be any different?"
[213] Green, "Legal Obligation and Authority".

than merely received them. But it is very unclear that the obligation generated by accepting these benefits would be a duty to obey. Suppose I accepted financial help from my friend. This may generate a duty on me to reciprocate him financially. But why does that generate a duty to obey?[214]

How about the benefit of order and security? I believe that even if we focus on these benefits, it is unclear why the reciprocal duty generated by accepting such benefits would be a duty to *obey* the law. If a rich person donates money to support an education campaign to preach the importance of law-abidingness, that may result in a greater improvement in order and security for other citizens than simply obeying the law himself.[215] But let us set this problem aside for there is a more obvious problem in justifying the claim that non-socially deprived people have a duty to obey the law by appealing to the idea that they voluntarily accepted the benefit of order and security. The problem is that such benefits are not reasonably avoidable for most citizens, *even the non-socially deprived ones*. It is hard to see how people can avoid the benefits of order and security provided by the state without migrating to some state-less area and even if they can pay for it, this is not a reasonable alternative for

---

[214] Simmons, "The Obligations of Citizens"; Simmons, *Moral Principles and Political Obligations,* 185.

[215] For another theory that falls prey to this objection, see Wellman, "Toward a Liberal Theory of Political Obligation", 749. Wellman defends a duty to obey the law by appealing to one's duty to do his fair share to maintain "political stability". One may ask: why is obeying the law necessary, or the most effective means, to do one's fair share to maintain political stability?

most of them given the social and personal costs.[216]

Since receipt and acceptance of benefit are not plausible grounds for justifying a duty to obey the law on the non-socially deprived offenders, even if socially deprived offenders have received or accepted fewer benefits from the state it does not follow that they would have a reduced duty to obey the law. Accordingly, I think we should reject the empirical claim of the general duty to obey the law argument.

Besides the success of the argument, I want to raise one more worry about the argument: I wonder if any but a very small amount of reduction of sentence can be justified by the general duty to obey the law argument, even if we assume that the argument is successful. This is because the *primary* wrongdoings in central mala in se, like murder or rape, must be the violation of the natural duties (the duty not to kill and the duty not to rape, etc.), which exist independent of the law, but not the violation of the content-independent general duty to obey the law.[217] As such, the difference in wrongfulness between a murder by a murderer who has a general duty to obey the law and a murder by a murderer who does not is minimal, and accordingly the difference in punishment of the two murderers justified by this difference in wrongfulness must

[216] Green, "Legal Obligation and Authority".
[217] Duff, *Punishment, Communication, and Community*, 56-59, 64-66, 182.

also be minimal.

4. Conclusion

In this chapter, I have examined two arguments for claiming that there is a partial justification that is more applicable to socially deprived offenders. According to the first argument, the no violation of natural duties argument, the proportion of socially deprived offenders who did not violate natural duties in committing their crimes is greater, and if an offender did not violate natural duties in committing his crime then his crime is less wrongful than the crime by an otherwise similar offender whose crime involved violation of natural duties. I argue that this argument is plausible, because socially derived persons have no natural duties to refrain from committing purely property crimes against rich people.

According to the second argument, the general duty to obey the law argument, the proportion of socially deprived offenders who had a reduced general duty to obey the law is greater, and if an offender had a reduced general duty to obey the law then his crime is less wrongful than the crime by an otherwise similar offender who had an uncompromised general duty to obey the law. I argue that this argument is implausible as it is not true that socially deprived offenders are more likely to have a reduced general duty to obey the law. In addition, I argued that even if the general duty to obey the law argument is successful, it can justify at most a very small amount of reduction of sentence.

Part III: Non-exculpatory Communicative Arguments

Chapter 7: Victims of the polity's injustice

1. Introduction

In part 3 of my thesis, which consists only of this chapter, I wish to focus on a non-exculpatory communicative argument, i.e. an argument that claims that the *right to communicate blame* to a socially deprived offender held by our courts and other penal institutions (for simplicity, when I simply refer to "courts" below I should be understood as referring to "courts and other penal institutions") is often compromised, and accordingly there is often a reason to punish a socially deprived offender less for his crime even if he is not less blameworthy for his crime.

Recently, many writers observe that in order to properly communicate blame to a wrongdoer for his wrongdoing, it is *insufficient* that the wrongdoer is blameworthy for the act. *In addition*, the *blamer* must have the standing or the right to communicate blame to the wrongdoer.[218] Drawing on this idea, Duff argues that, given that punishment involves communication of blame, if our courts' standing to communicate blame to socially deprived offenders is compromised, then punishment of those offenders is particularly problematic even if they are equally blameworthy for their crimes. According to Duff, our courts' standing to communicate blame to socially deprived offenders is compromised because those offenders, unlike non-socially deprived offenders, are victims of the polity's injustice (i.e. they suffered because the polity acted unjustly towards them).[219] Let me lay out the two main claims of this argument, which can be called the *polity's injustice argument*, before assessing it:

The Polity's Injustice Argument

Empirical claim: The proportion of socially deprived offenders who are victims of the polity's injustice (i.e. they suffered because the polity acted unjustly towards them) is greater than that for those who are non-socially deprived offenders.

Moral claim: If an offender is a victim of the polity's injustice, then our courts'

---

[218] Cohen, "Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists?"; Smith, "On Being Responsible and Holding Responsible".

[219] Duff, "Principle and Contradiction in the Criminal Law"; Duff, "Whose Law is it Anyway? Inclusion, Exclusion and the Criminal Law"; Duff, *Punishment, Communication, and Community*, 179-201; Duff, *Answering for Crime*, 191-194; Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial". This argument is also hinted to by Bazelon, but Duff is the first one, to my knowledge, who provides an extended defence of the argument. See Bazelon, *Questioning Authority.* See also Garvey, "Injustice, Authority, and the Criminal Law".

standing to communicate blame to him is compromised, and accordingly there is a reason to punish him less, than an otherwise similar offender who is not a victim of the polity's injustice.

The remaining parts of this chapter will be arranged as follows. In section 2, I will assess the empirical claim of the argument. I will argue that the empirical claim is true, on a particular understanding of the concept of "polity". In section 3, I will discuss the moral claim of the argument. I will argue that the moral claim of the polity's injustice argument is unmotivated by attacking Duff's argument for it.

2. The empirical claim of the polity's injustice argument

I think it is undeniable that many socially deprived offenders, unlike non-socially deprived offenders, suffer from a lack of important social, economic, and political goods. But is it true that their disadvantage is a result of the fact that "the polity" acted unjustly towards them?

I believe it is true on one interpretation of the idea of "polity". Let us distinguish two things Duff can mean by "polity" and accordingly two things Duff can mean by saying that the polity acted unjustly towards socially deprived offenders.[220] On the first interpretation, to say that the polity acted unjustly means that *all* individual members of the polity acted unjustly. On this interpretation, the empirical claim is false: there are saints and heroes who fight for distributive justice and sacrifice much for the poor in every society.

On the second interpretation, to say that the polity acted unjustly means that the polity, *as a collective*, acted unjustly.[221] We should notice that by saying that a collective did something we do *not* imply that everyone constituting the collective did that thing: by saying that the US went to war we do not imply that every US citizen went to war. And I think on this interpretation, the empirical claim is plausible. The disadvantage suffered by most socially deprived offenders is not an unalterable natural fact – as the polity or the society, as a collective, has much power to eliminate the disadvantages of the socially deprived people (at least assuming that the political structure is largely democratic). The failure to do so can be regarded to be an act of injustice on behalf of the polity *as a collective*. Since the empirical claim is true only on the second

[220] Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial", 137-138. Similar phrases can be found in Duff, *Punishment, Communication, and Community*, 186-187, 196-197. Sometimes Duff does not use the word "polity" but rather the phrase "political community". I take the two terms to stand for similar ideas.

[221] This interpretation can find textual support in Duff's work, as Duff sometimes talks about "the collective hand" of the polity, etc. See Duff, *Punishment, Communication, and Community,* 188; Duff, *Answering for Crime,* 191-92.

understanding of the concept of polity, I will understand the concept of polity in the second sense in the rest of this chapter.

3. The moral claim of the polity's injustice argument

The moral claim of the polity's injustice argument has two parts. The first part is the idea that if our courts' standing to communicate blame to an offender is compromised, then there is a reason to punish him less. The second part is the idea that if an offender is a victim of the polity's injustice, then our courts' standing to communicate blame to him is compromised as compared to their standing to communicate blame to an offender who is not a victim of the polity's injustice.

I do not want to question the first part of the moral claim of the polity's injustice argument. Instead, I want to question the second part, the idea that if an offender is a victim of the polity's injustice, then our courts' standing to communicate blame to him is compromised (as compared to their standing to communicate blame to an offender who is not a victim of the polity's injustice). I will discuss Duff's main argument for the idea, which I will call *the polity's standing to communicate blame argument,* and I will argue that it is unsound.

3.1. The polity's standing to communicate blame argument

The main argument given by Duff for the idea that if an offender is a victim of the polity's injustice, then our courts' standing to communicate blame to him is compromised (as compared to their standing to communicate blame to an offender who is not a victim of the polity's injustice), *the polity's standing to communicate blame argument*, consists of two premises:[222]

P1: If an offender is a victim of the polity's injustice, then our *polity's* standing to communicate blame to him is compromised as compared to its standing to communicate blame to another offender who is not a victim of the polity's injustice.

P2: If our *polity's* standing to communicate blame to an offender is compromised, then our *courts'* standing to communicate blame to him is compromised as well.

I will discuss P1 in section 3.1.1 and P2 in section 3.1.2. I will argue that while some arguments offered for P1 in the literature are unsound, there is a sound argument for

[222] Duff, *Punishment, Communication, and Community*, 185-186.

P1. However, I will argue that we should reject P2.

Before I go on, I wish to clarify what Duff means by claiming that the polity's standing to communicate blame to an offender is compromised. Does he mean that the individual right to communicate blame to the offender is compromised, *for everyone constituting the polity*? Duff certainly does not mean that as he recognized that, for example, the victims may still retain their individual rights to communicate blame to an offender even if the polity lost its right to communicate blame to him.[223] So what does Duff mean there? I think what he means is that the right held by the polity, *as a collective*, to communicate blame to such an offender is compromised. Just like saying that the polity, *as a collective*, acted unjustly does *not* mean that *everyone* in the polity acted unjustly (as we have seen above in section 2 of this chapter), analogously, that the right to communicate blame to such an offender held by the polity, as a collective, is compromised does not entail that, for every member of that polity, his individual right to communicate blame to the offender is compromised.

In order for this interpretation of the phrase "the polity's right to communicate blame to offenders who are victims of its injustice is compromised" to make conceptual

[223] Duff, *Punishment, Communication, and Community*, 188.

sense, it is necessary that collective entities, like polities, can hold rights (and as a result, can lose rights). Some people may be skeptical as to whether collective entities can be right holders.[224] To a certain extent I am sympathetic with that view but I wish to make two points. First, if it does not make sense to say that collective entities can hold rights and lose rights, then it is a problem for Duff but not for me who intends to criticize his argument. Second, I want to argue that *even if* we agree with Duff that collective entities, like polities, *can* hold rights and thus can lose rights, the polity's standing to communicate blame argument is still unsound.

3.1.1. Is the polity's standing to communicate blame to victims of its injustice more compromised than its standing to communicate blame to other offenders?

In section 3.1.1, I will explore three arguments for P1 of the polity's standing to communicate blame argument. What is the basis for believing that the polity's standing to communicate blame to offenders who are victims of its injustice is more compromised than its standing to communicate blame to other offenders? I will argue against two arguments for P1 (the hypocrisy argument and the complicity argument) but I will argue that one argument for P1 (the reciprocity argument) is sound.

[224] Some of related issues are explored in Jones, "Group Rights".

Before I assess the three arguments for P1, let me first clarify that the issue is *not simply* whether the polity's standing to communicate blame to offenders who are victims of its injustice is compromised. Rather, the issue is whether the polity's standing to communicate blame to the victims of its injustice is *more compromised than its standing to communicate blame to other offenders*.[225] This is because if the polity's standing to communicate blame to both groups of offenders is equally compromised, then we cannot draw a distinction between victims of the polity's injustice and other offenders, which is necessary for the moral claim of the polity's injustice argument.

3.1.1.1. Hypocrisy

The first argument for the claim that our polity's standing to communicate blame to victims of its injustice is compromised in a way that its standing to communicate blame to other offenders is not, the *hypocrisy argument*, claims that hypocrisy undermines a blamer's standing to communicate blame, i.e. if you committed an equally or more serious wrong than an offender without accounting for your own

[225] Duff, "Principle and Contradiction in the Criminal Law", 196; Duff, *Punishment, Communication, and Community*, 195, 197; Duff, *Answering for Crime*, 191-93.

wrong, then your standing to communicate blame to him is compromised; since our polity committed serious wrongs without accounting for them, it is hypocritical for it to communicate blame to offenders who are victims of its injustice.

Some critics of Duff focus on attacking this argument,[226] though I believe that there are good textual reasons to believe that Duff does not intend to rely on hypocrisy in explaining why our polity's standing to communicate blame to offenders who are victims of its injustice is compromised.[227] Nevertheless, this argument is interesting in its own right and merits discussion whether or not it was offered by Duff.

Many people have the intuition that hypocrisy undermines a blamer's standing to communicate blame: The pot should not call the kettle black. But I doubt if this intuition is universally shared. I have heard, from time to time, that pots say, "but that is another business. We are talking about the kettle's blackness at this moment."[228] So can the claim that hypocrisy undermines a blamer's standing to communicate blame be justified on theoretical grounds?

[226] Holroyd, "Punishment and Justice"; Matravers, "Who's Still Standing: A Comment on Antony Duff's Precondition of Criminal Liability".
[227] It is clearest when we read his recent piece "Blame, Moral Standing, and the Legitimacy of the Criminal Trial".
[228] Melissaris, "Property Crimes as Crimes of Injustice".

I think Victor Tadros and Jay Wallace (independently, though their arguments are similar for my purpose here) have offered a compelling justification for the claim that hypocrisy undermines a blamer's standing to communicate blame.[229] Their justification for the claim runs as follows: everyone has an interest in not being censured (by "censuring" someone I mean communicating blame to him). If I communicate blame to someone for a certain fault, without holding myself accountable for similar faults, then my pair of conducts implies that I treat my interest not to be censured to be more important than the offender's interest not to be censured, and thus, violates the ideal of *moral equality*.

I think the ideal of moral equality provides a solid justification for the claim that hypocrisy undermines a blamer's standing to communicate blame. But even so I think the hypocrisy argument cannot justify P1 of the polity's standing to communicate blame argument. It is because, as argued by Tadros and Holroyd, the idea of hypocrisy cannot distinguish between wrongdoers who suffered in the hands of the blamer and other wrongdoers and hence, cannot justify the claim that our polity's standing to communicate blame to offenders who are victims of its injustice is compromised *in a way that* its standing to communicate blame to other offenders is

[229] Wallace, "Hypocrisy, Moral Address, and the Equal Standing of Persons", 324, 328-29; Tadros, "Poverty and Criminal Responsibility", 401-404.

not.[230] Hypocrisy forbids a robber from communicating blame to *any* petty thief, *whether or not the petty thief was the victim of his robbery*. Similarly, even if the idea of hypocrisy can explain why our polity's standing to communicate blame is compromised because of its past injustice, it cannot explain why our polity's standing to communicate blame to offenders who are victims of its injustice is *more* compromised than its standing to communicate blame to other offenders.

3.1.1.2. Complicity

The second argument for the claim that the polity's standing to communicate blame to offenders who are victims of its injustice is compromised in a way that its standing to communicate blame to other offenders is not relies on the idea of *complicity*. Before I go on, let me state that this argument, while quite relevant for our purpose here, is not one given, or endorsed, by Duff.[231] It is an argument given by Tadros.[232]

The complicity argument consists of two claims. The first claim is that complicity is an *independent ground* for compromise of standing, that is, it leads to a compromise

230 Holroyd, "Justice and Punishment", 84, 89; Tadros, "Poverty and Criminal Responsibility", 409.
231 Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial", 137.
232 Tadros, "Poverty and Criminal Responsibility", 393.

of standing to communicate blame *over and above the charge of hypocrisy*.[233] To illustrate this claim, imagine two offenders, Rachel and Tom, who committed two identical crimes independently. Paul is an accomplice of Rachel, but he is *not* an accomplice of Tom. According to the first claim of the complicity argument, Paul's standing to communicate blame to Rachel is more compromised than his standing to communicate blame to Tom.[234] (If hypocrisy is the only relevant charge here, then Paul's standing to communicate blame to Rachel and his standing to communicate blame to Tom should be equally compromised – since he would be acting hypocritically no matter he is blaming Rachel or Tom.) The second claim of the complicity argument is that our polity is an accomplice of the crimes committed by victims of its injustice but not an accomplice of the crimes committed by other offenders, because our polity wrongfully deprives victims of its injustice important goods with the knowledge that the lack of such goods is criminogenic.[235]

I have some reservations about the second claim of the complicity argument. The main problem is that affluence may be the cause of the crime for some offenders. Just think about some spoiled young offenders or some offenders who committed white

[233] Tadros clearly believes that the two are distinct grounds for compromise of standing to communicate blame. See Tadros, "Poverty and Criminal Responsibility", 399-400.
[234] The example is directly taken from Tadros, "Poverty and Criminal Responsibility", 404.
[235] Tadros, "Poverty and Criminal Responsibility", 404-409.

collar crimes. They certainly are not *victims* of the polity's injustice. However, given that their affluence is also caused by the polity's injustice, the polity would be an accomplice of their crimes, in addition to being an accomplice of the crimes committed by victims of its injustice. So it is unclear whether the complicity argument can draw a sharp distinction between offenders who are victims of the polity's injustice and other offenders.

But I want to focus on attacking the first claim of the complicity argument. Why should we accept that complicity is an independent ground for compromise of standing? I do not have a strong intuition that Paul's standing to communicate blame to Rachel is more compromised than his standing to communicate blame to Tom. As we have seen above in section 3.1.1.1, there is a strong theoretical justification, based on moral equality, for claiming that hypocrisy undermines standing. But I cannot see any good theoretical justification for the claim that complicity is an independent ground for compromise of standing over and above the charge of hypocrisy. I will examine the proposed justification for that claim offered by Tadros, and I will argue that it is unsound.

Tadros tries to explain why "claims based on complicity erode standing to a greater degree than claims based on hypocrisy" in the following passage:

> Let us suppose that Paul engages Rachel in a practice of responsibility about a particular wrongful act which he is complicit in, but he blocks attempts to hold him responsible for that act. In doing so, he implies that Rachel must respond to her wrongdoing by recognizing that the interest of the victim underpins a moral principle that she ought not to have perpetuated the wrong. But he denies that he must respond to the interest in that way and through that repair his relationship with the victim. The victim is provided with an opportunity for the individuals responsible to repair their relationship with her by accepting the moral significance of her interests. In holding Rachel responsible for the wrong while denying his responsibility, Paul implies that the moral obligation to repair applies especially to her with respect to this instance of wrongdoing. Hence, the claim that he and Rachel face different moral demands is especially stark. In cases of hypocrisy, [i.e. when Paul is communicating blame to Tom], in contrast, Paul might claim that others are entitled to hold him responsible for the wrongs he has perpetuated, where the wrongs have set back the interests of victims."[236]

[236] Tadros, "Poverty and Criminal Responsibility", 404.

I find this passage a bit puzzling, but Tadros's point seems to be that Paul can say to Tom "I will account to my victim (i.e. the victim of the crime committed by Rachel, of which Paul is an accomplice) in another occasion" as a defence of his standing when he is communicating blame to Tom. However, Paul cannot raise that defence when he is communicating blame to Rachel. Since this defence of his standing is available when he is communicating blame to Tom but unavailable when he is communicating blame to Rachel, his standing to communicate blame to Rachel is more compromised than his standing to communicate blame to Tom.

In reply there are two problems with this justification. First, why can't Paul raise that defence when he is communicating blame to Rachel? Why can't he say "I will account to our victim in another occasion" when Rachel challenges his standing? Second, why can Paul legitimately raise that defence when he is communicating blame to Tom if we know that Paul is not going to account to his victims (*if* he is really going to account to them *then he is not even a hypocrite*)?

Now Tadros may reply that I have overlooked the point that our practices of responsibility, including communication of blame, are "act focused". He writes,

> "To answer [the question of why complicity erodes standing over and above hypocrisy] we must explore a further feature of our practices of responsibility: their focus on particular acts. Practices of responsibility have value in repairing our moral relations with each other. They primarily achieve this by being act focused. They are focused on holding people responsible for the particular acts that they perform rather than for their moral conduct in general."[237]

His explicit definition of what is meant by saying that practices of responsibility are "act focused" is, as seen in the quotation above, that "they are focused on holding people responsible for the particular acts that they perform rather than their moral conduct in general." But this does not seem to help Tadros to draw a distinction between communicating blame to Rachel and communicating blame to Tom. No matter whether Paul is communicating blame to Tom or to Rachel, the process can be act-focused – it can focus on the particular wrong acts by Tom and Rachel, not the general moral conduct of them.

[237] Tadros, "Poverty and Criminal Responsibility", 404.

My critique of the claim that complicity erodes standing over and above hypocrisy will be more satisfying if I can offer a diagnosis of why some people find the claim persuasive. My diagnosis is that many people who accept that claim make two mistakes: they confuse compromise of standing to communicate blame with *conversational awkwardness*; and they only focus on *one special type of complicity cases*. Let me explain. If I participated in a wrongdoing (say, murder) and *shared the aim* of my co-offenders, then it would be more *conversationally awkward* for me to criticize my co-offenders than to criticize other murderers because my action signifies approval of *the precise act* of my co-offenders in a way that does not apply to other murders. But this only establishes conversational awkwardness but not a compromise of standing.[238] (Not every awkward action is a violation of right.) And in any case, the above situation is one in which the accomplice *shares the aim* of the principal, which is a special case of complicity. Given that our polity, even if unjust, is unlikely to share the aims of the offenders who are victims of its injustice,[239] this special case cannot be generalised to establish the claim that the charge of complicity erodes our polity's standing to communicate blame to offenders who are victims of its injustice

[238] Wallace also draws a distinction between conversational awkwardness and a loss of right to communicate blame. See Wallace, "Hypocrisy, Moral Address, and the Equal Standing of Persons", 321.
[239] Tadros himself notes this point. See Tadros, "Poverty and Criminal Responsibility", 405.

over and above the charge of hypocrisy.

3.1.1.3. Reciprocity

The third argument for P1 of the polity's standing to communicate blame argument, the claim that the polity's standing to communicate blame to offenders who are victims of its injustice is compromised in a way that its standing to communicate blame to other offenders is not, that I will consider is based on the idea of *reciprocity*. This is the argument that Duff focuses on in his more recent works,[240] and I also find this argument the most plausible amongst the three. (It is a curious fact that most critics of Duff in the literature do not focus on this argument but rather on the hypocrisy argument. So I believe most critics failed to engage with the strongest argument by Duff for the claim that the polity has lost its right to communicate blame to offenders who are victims of its injustice.)

The reciprocity argument consists of the following two claims:

(a) in order to have the (uncompromised) right to communicate blame to an offender

[240] See, e.g., Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial".

in criminal punishment, the blamer must have treated the offender as a citizen; and

(b) Our polity failed to treat offenders who are victims of its injustice as citizens.

I think (b) is uncontroversial, so let us focus on (a). Why is a blamer's right to communicate blame to an offender in criminal punishment conditional upon the fact that the blamer has treated the offender as a citizen? This question is a legitimate one, once we notice that *not all rights are conditional*. For example, our duty to refrain from raping seems to be unconditional: we have a duty not to rape rapists.[241] So Duff needs to explain why the right to communicate blame in criminal punishment is a kind of right that is conditional, depending on whether the blamer treated the offender as a citizen, unlike the right not to be raped.

I think Duff can defend (a) based on two further claims. (The first further claim was explicitly made by Duff. The second has not been explicitly offered by Duff, but I think Duff would accept it as he accepted something similar.) The first further claim for (a) is that an offender's duty to account in criminal law is not owed towards everyone but *only owed towards his fellow citizens*.[242] Duff gives two independent

[241] Scanlon, *Moral Dimensions*, 139-145.
[242] Duff, *Answering for Crime*, 48.

reasons for this further claim. The first reason for the first further claim is that ordinary criminal wrongs (i.e. wrongs that do not count as international crimes, crimes against humanity, etc.) are wrongs that concern only citizens, not all human beings. Duff thinks that, for example, a theft committed by a Pole against a Pole in Poland is not the business of English citizens.[243] The second reason for the first further claim is that whether or not criminal wrongs are only accountable to citizens, blame communicated in the criminal process is a kind of blame that addresses the offender as a citizen, and as such, it seems that the offender only has a duty to listen or account if the blamer is one of, or a group of, his fellow citizens.[244] The second further claim for (a) is that for duties that we have only in virtue of being members of thick moral communities (like citizenship), we owe them only towards members of that community who have discharged their duties towards us, unlike duties that we have simply in virtue of being human.[245] This further claim, I think, is also plausible. It seems plausible to think that if my student does not treat me as his teacher, then he has no right to insist me to continue to treat him as my student. The distinction between the duties we have in virtue of being a member of a thick moral community

---

[243] Duff, "Authority and Responsibility in International Criminal Law", 595.

[244] Duff, *Answering for Crime*, 49-51; Duff, "Blame, Moral Standing, and the Legitimacy of the Criminal Trial", 138

[245] Duff explicitly accepted that duties of citizenship are reciprocal. See Duff, "Whose Law is it Anyway?", 6. But he has not explicitly accepted my explanation as to why duties of citizenship are reciprocal (my explanation is that they are duties generated by membership in thick moral communities), though I think my explanation would be acceptable to him.

and the duties we have simply in virtue of being human helps to explain why the right not to be raped is not conditional on reciprocity but the right to communicate blame in criminal punishment is.

So I believe the challenge to the reciprocity argument can be adequately met, and thus we should accept P1 of the polity's standing to communicate blame argument.

3.1.2. The relationship between the polity's standing to communicate blame and the courts' standing to communicate blame

Now let us move on to P2 of the polity's standing to communicate blame argument. Why should we accept that if our *polity's* standing to communicate blame to an offender is compromised, then our *courts'* standing to communicate blame to that offender would be compromised as well?

Duff seems to believe that is so because our courts have the right to communicate blame to an offender only if the persons whom the courts represent have the right to communicate blame to that offender, and that our courts *must* be seen as acting as *the representatives of*, or to use the exact phrases by Duff, *"in the name of", or "on* *behalf of",* the whole polity.[246]

I want to argue, to the contrary, that our courts need *not* be seen as acting on behalf of the whole polity in criminal proceedings. They can be seen as acting on behalf of a *subset* of the citizens, namely, *the just citizens*, in criminal proceedings. Since the just citizens, who have not committed injustice towards offenders who are victims of the polity's injustice, retain their standing to communicate blame to those offenders despite the fact that our polity is unjust, the fact that our polity has compromised its standing to communicate blame to offenders who are victims of its injustice does not entail that our courts have compromised their standing to communicate blame to those offenders. Let me call this objection against Duff *the different representee objection.*

In fact, I would think this is a natural response. Perhaps an analogy can illustrate my point. Suppose that I promised both of my parents to do a task, say, to attend a university. Now, for some reason, my father loses his right to enforce his promise. Perhaps, for example, he waived his right to enforce my promise and thus released me from my duty to him. Suppose now my sister is sent as a representative by my parents to ask me to attend a university. If I reply to her that I no longer owe the duty to my

[246] Duff, *Punishment, Communication, and Community,* 185-188; Duff, *Answering for Crime,* 53.

parents as a whole, it seems that she can legitimately reply that her standing to enforce the promise is not compromised. She can say something like: yes, you may not owe a duty to our parents as a whole now, but you still owe a duty to our mother, and I am now acting on behalf of her. Analogously, why can't the courts say, in response to a challenge to their standing to communicate blame by offenders who are the victims of the polity's injustice, the following: "yes, you may not owe a duty to account for your crimes to the polity as a whole now, but you still owe a duty to account to the just citizens, and I am now calling you to account on behalf of them"?

Duff has not explicitly considered this attack on P2 of the polity's standing to communicate blame argument so it is hard to know what his exact reply would be. Without implying that he would accept any of the following, I wish to outline six objections to my critique on behalf of him. I will argue that all of them fail.

The first objection can be called the *authorisation objection*. According to it, since the just citizens have not authorised the courts to act on behalf of them, the courts cannot be seen, as my position would hold, as acting on behalf of them.

I have two replies to this objection. My first reply to the authorisation objection is that if authorisation is the problem, then the problem can easily be solved by asking the just citizens to authorise the courts. The standing of the courts can then be recovered overnight, if just citizens do whatever is required (e.g., submitting letters) to authorise them. In other words, as long as *some* of the just citizens want the offenders who are victims of the polity's injustice to be punished in our criminal courts, the courts' standing can easily be recovered. Duff's argument then fails to show that just punishment of offenders who are victims of the polity's injustice is impossible in our actual society, as radical restructuring of the society towards social justice in the ways that Duff envisaged is unnecessary for the courts to recover their right to communicate blame to those offenders.[247]

My second reply to the authorisation objection is that if lack of authorisation is a problem, the problem is a general one that applies to punishment of all offenders, whether they are victims of the polity's injustice or not, and thus cannot show that there is something particularly problematic about punishment of offenders who are the victims of the polity's injustice. To illustrate, suppose we insist that authorisation requires actual explicit consent. On this test, of course, the just citizens have not authorised the courts to punish the offenders who are victims of the polity's injustice

[247] Duff has discussed what kinds of social reform are required to restore the legitimacy of punishing socially deprived offenders. See Duff, *Punishment, Communication, and Community*, 200–201; Duff, *Answering for Crime*, 192–193.

in most societies. But on this test it is hard to see how the courts have been authorised by the whole polity, as Duff would want to say, to punish offenders who are *not* victims of the polity's injustice in most societies as well.[248]

The second objection can be called the *control objection.* According to it, since the just citizens do not control the criminal proceedings, our courts cannot be seen as acting as the representative of them.[249] Notice that this objection is different from the authorisation objection, since the concept of control is different from the concept of authorisation.[250] I can authorise something without having control over whether it will happen or not. For example, suppose that I authorised my agent to do X. However, my agent would have done X whether or not I have authorised him. In this case, I have authorised X to be done, but I have no control over whether X would be done.

In reply to this objection, it is perhaps true that the polity as a whole has control over criminal proceedings in a way that the just citizens, who may be only a minority of

---

[248] It is worthwhile to notice that Duff, in the context of discussing the legitimacy of international criminal law, recognises that in order for a court to act for some people, what is needed is only "some normative relationship" between the court and those people and actual explicit consent is not necessary: Duff, "Authority and Responsibility in Criminal Law", 599.

[249] Haque seems to endorse the idea that authorisation requires control, as he thinks that in order to for a criminal process to act for the victims they must have "influence" over the process. See Haque, "Revolution and the Criminal Law".

[250] For the idea of control over a legal process, see Duff, *Punishment, Communication, and Community*, 61.

the population, do not, at least in a broadly democratic state. But it seems to me that a representative can legitimately act for his representee as long as he is authorised by the representee, and control is not a necessary condition. The case of the agent above aptly illustrates this point. Suppose that I authorised my agent to censure a wrongdoer on behalf of me and I am the only person in the world who has the right to censure the wrongdoer. Suppose, however, that my agent hated the wrongdoer so much that even if he was not authorised by me he would still censure that wrongdoer. Even so, it seems to me that my agent can legitimately say he is acting on behalf of me in censuring the wrongdoer even though I do not have control.

Moreover, we should notice that we sometimes think that we can legitimately act for someone even if he does not have control over the process. Think about how we sometimes say that we act as trustees or representatives for mentally disabled persons or children without them having much control over the process.

The third objection can be called the law of the *whole polity objection*. According to this objection, the courts must, in a criminal process, communicate the voice of the law; but the content of our criminal law (at least for statutes in a democratic polity) depends on the whole polity; therefore, courts must communicate on behalf of the whole polity in criminal punishment.[251]

In order to assess this objection, we need to clarify one point. What exactly is meant by saying that we communicate blame to someone in the "voice of the law"? This phrase is metaphorical. A natural interpretation of the phrase is that we are communicating blame to him *because* he violated the law; the object of blame is violation of the law. So understood, this objection can be broken down into the following points:

a) The blame communicated in criminal punishment is a kind of blame for violation of the law.

b) The content of our laws depends on the polity as a whole.

c) If our courts are communicating blame to someone for violation of the law, then they must be communicating blame on behalf of the whole polity.

d) Therefore, courts must be communicating blame on behalf of the whole polity in criminal punishment.

This objection, as interpreted, suffers from two problems. First, I think (c) is doubtful.

[251] I am grateful to Antony Duff for suggesting this objection.

The underlying assumption of (c) seems to be that only the entity on whom the content of some rules depends can communicate blame to a person for violation of those rules. But that assumption seems to be false. We can communicate blame to a football player for violating some football rules even if the content of football rules does not depend on us. So analogously why can't the courts act on behalf of the just citizens in communicating blame to an offender for violating the criminal law, even if the content of our criminal law does not depend (only) on the just citizens?

Second, and more importantly, I think (a) is false (and I believe Duff would agree with me on this point). Blame communicated in criminal punishment is, as Duff himself has pointed out, primarily a kind of blame, not for violation of the law, but for violation of duties we are accountable towards our fellow citizens, duties that can exist independently of the law: when we punish a murderer, the primary condemning message we should send is not against rule-breaking.[252] Rather, we should be blaming the offender for murdering, a kind of wrong that can be committed even if there are no legal rules.

I want to reply to a rejoinder by my opponent before moving on to the fourth

---

[252] Duff himself does not seem to dispute this point. He claims that it is mistaken to see the violation of the law, *per se*, as the object of blame, at least for the central mala in se. See Duff, *Trials and Punishments*, 211–217.

objection. My opponent may argue: if the object of blame in criminal punishment is the violation of duties accountable to citizens *per se* (which exist independently of the law) but not the violation of the law, as I have claimed above in arguing for the falsity of (a), does it mean that the question of whether a person violated a valid criminal law is irrelevant to criminal punishment? Does it entail some kind of extreme natural law view in which a person could be legitimately punished without violating any criminal law (as long as he violated some natural duties that are accountable to his citizens)? Not at all. Borrowing a distinction from the work of Duff, to say that the violation of the law should not be the *object* of blame in the criminal process does not mean that the violation of the law should not be a *condition* of blame in the criminal process.[253] There are good reasons, like fair warning and the rule of law, to take violation of the law to be a condition of the criminal process even if violation of the law is not the object of blame in the criminal process.

The fourth objection can be called the *diminution of standing objection.*[254] According to it, the *just citizens' right to communicate blame to* an offender is *less* than the *whole polity's right to communicate blame to* that offender. Therefore, since the courts can act for the whole polity in communicating blame to offenders who are not

[253] Duff, "Virtue, Vice, and Criminal Liability: Do We Want an Aristotelian Criminal Law?", 155.
[254] I am grateful to Massimo Renzo for suggesting this objection.

victims of the polity's injustice but the courts can only act for the just citizens in communicating blame to offenders who are victims of the polity's injustice, the courts' standing to communicate blame to offenders who are victims of the polity's injustice is compromised as compared to their standing to communicate blame to other offenders.

But why should we believe that the polity's right to communicate blame is stronger than the right held by the just citizens? One suggestion would be that *number* matters: since the number of people constituting the whole polity must be larger than the number of just citizens, the polity's right to communicate blame is stronger. But this suggestion will face counter-intuitive implications: does it mean, for example, that courts in China will have a stronger right to communicate blame to an offender than the courts in England, since China has a larger population? While the strength of some rights depends on numbers, it is clear that the strength of other rights, including the right to communicate blame, does not.

Another suggestion for why the polity's right to communicate blame is stronger than the just citizens' right to communicate blame is that the polity's right to communicate blame is an irreducibly collective right. On this line of thought, the right to communicate blame held by the polity is an irreducibly collective right, entailing a duty on the part of the offender to account, not merely towards all individual citizens, but also towards the polity as an irreducible collective; it is more than the aggregate of the individual rights to communicate blame held by all citizens constituting the polity, and thus it is more than the aggregate of the individual rights to communicate blame held by the just citizens.

To this suggestion my first reply is that it is unclear that there are really irreducibly collective rights. My second reply is that even if we can make sense of the idea of irreducibly collective rights, the particular collective right on the table is extremely obscure. For that right implies a duty to account to a collective entity, which cannot be reduced to a duty to account to individual persons constituting the entity. I can understand why, after committing some wrong, I ought to account to some individual persons, like victims or other citizens. I need to repair my relationship with them. But it is wholly mysterious to me why, after committing some wrong, I ought to account to some non-personal collective entity.

The fifth objection can be called the *public resources objection*.[255] According to this

[255] I am grateful to Victor Tadros for suggesting this objection.

objection, public resources can only be legitimately used for the pursuit of aims on behalf of the whole polity. If courts in criminal proceedings act on behalf of only a subset of citizens but use resources that belong to the whole polity, then they act illegitimately.

I have two replies to this objection. My first reply to the public resources objection is that even if this objection can successfully establish the claim that the punishment of offenders who are victims of the polity's injustice is illegitimate, we should observe that *the kind* of illegitimacy involved is very different from the kind of illegitimacy originally alleged in Duff's argument. According to Duff's argument, punishment of offenders who are victims of the polity's injustice is illegitimate because it violates *their* rights. On the public resources objection, in contrast, the illegitimacy of punishing the offenders who are victims of the polity's injustice lies in the violation of rights, *not of the* offenders who are victims of the polity's injustice*, but of the unjust citizens,* since they are the ones who can complain that the courts are not acting on behalf of them while using their resources.

My second reply to the public resources objection is that as (a) the courts do not violate the right of a person if they use his resource to do something that he has a duty to do even if the courts do not act in his name in doing so; and (b) unjust citizens have a duty to communicate blame to offenders or assist in the process that enables blame to be communicated to offenders, it follows that the courts can legitimately use the unjust citizens' resources to fund a punitive system even if the system is not acting in the name of the unjust citizens. I think (a) is relatively uncontroversial. I think (b) is true as well. It will be beyond the scope of this thesis to defend the claim that every citizen has a duty to communicate blame to offenders or assist in the process that enables blame to be communicated to offenders. But it is worthwhile to disarm one objection against (b). My opponents may object to (b) by saying that since unjust citizens lose their right to communicate blame to an offender, they also lose their duty to communicate blame to that offender. But this objection to (b) is unsound, since it does not generally follow from the fact that I lose the right to do X that I also lose the duty to do X. The fact that I lose the right to do X normally means, instead of an evaporation of the duty to do X, that now I face a moral dilemma: I have a duty both to do X and not to do X.

The final objection can be called the *lynching objection*. According to this objection, if the courts do not act on behalf of the whole polity in the criminal process, then the process looks, disturbingly, like lynching.[256]

In reply, if the courts act for the just citizens, instead of the whole polity, in punishing offenders who are victims of the polity's injustice, then this process of course, trivially, shares one feature with lynching: namely, it is not a kind of punishment on behalf of the whole polity. But in order to understand whether it is therefore objectionable we need to first answer what is so problematic about lynching. I can think of several problems with lynching. First of all, lynch mobs sometimes act angrily and thus the terms of punishment inflicted by them are often disproportionate. Second, one who is a subject of lynching is usually not accorded the relevant procedural protections in trial. Third, lynching disrupts the rule of law, by suggesting people to break the law and take the law in their own hands. Fourth, sometimes a person who committed a wicked act without violating the law is lynched, and as such, lynching sometimes violates the requirement of fair warning.

But none of these problems applies (at least none of them necessarily applies) to my suggestion, which requires us to understand the courts as acting on behalf of the just citizens rather than the whole polity in punishing offenders who are victims of the

[256] I am grateful to Antony Duff and Victor Tadros for suggesting that I should think about this objection.

polity's injustice. First, I have only suggested a change in understanding who can the courts act on behalf of in a criminal process. I have not advocated any change in gravity of punishment. As such, punishment on behalf of just citizens need not be disproportionate. Second, in proposing that the courts can act on behalf of the just citizens rather than the whole polity, I have not advocated any change in trial procedure, and I cannot see any such change would be entailed if the courts are acting on behalf of the just citizens rather than the whole polity. As such, I can see no reason why defendants must be offered fewer procedural safeguards if the courts communicate blame to them on behalf of just citizens rather than on behalf of the whole polity. As to the third and the fourth problems, the courts need not be violating the rule of law when they punish an offender on behalf of the just citizens. The questions of whether violation of a valid law is a necessary condition of criminal punishment and whether and when a valid law should be obeyed are, so far as I can see, entirely different questions from the issue of who the courts are acting on behalf of in criminal punishment.

4. Conclusion

In this chapter I have examined the polity's injustice argument. According to this argument, socially deprived offenders are more likely to be a victim of the polity's injustice; and if an offender is a victim of the polity's injustice, then our courts' standing to communicate blame to him is compromised; and if our courts' standing to communicate blame to an offender is compromised then there is a reason to punish him less for his crime even if he is equally blameworthy for his crime.

I have argued that the empirical claim of the polity's injustice argument is true once we take the *collective* interpretation of the concept of "polity". I have, however, tried to reject the moral claim of the argument by arguing that the courts need not act on behalf of the whole polity in communicating blame in criminal punishment.

Part IV: Non-exculpatory Non-communicative Arguments

Chapter 8: Distributive justice and improvement of the worst off

1. Introduction

In part 4 of this thesis, which consists of this chapter and the next two chapters, I will examine four non-exculpatory non-communicative arguments (one in this chapter, one in chapter nine, and two in chapter ten), i.e. arguments for claiming that there is a ground to punish an offender less for his crime, that is based neither on the idea that the offender is less blameworthy for his crime nor the idea that our courts' standing to communicate blame to him is compromised, that is more applicable to socially deprived offenders.

The argument that I am going to focus on in this chapter, which can be called the *improvement of the worst off argument*, consists of the following two claims:

Improvement of the Worst Off Argument

The moral claim: If by reducing the punishment of an offender, we will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices), then there is a reason to reduce his punishment that does not apply to offenders the reduction of punishment of whom will not lead to that consequence; and accordingly there is a reason to punish him less for his crime than an otherwise similar offender the reduction of punishment of whom will not lead to that consequence.

The empirical claim: For a greater proportion of socially deprived offenders than non-socially deprived offenders, reducing their punishment will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices).

Many people reject the moral claim of the improvement of the worst off argument. They believe that the improvement of the well-being of the worst off members of society is not a proper concern of punishment.[257] I believe they are mistaken. I will argue that the moral claim is true because distributive justice, which requires us to

[257] See the writers referred to in section 2.4 of this chapter below.

improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices), provides reasons for action for sentencing institutions. In fact, I will argue that if mitigating the punishment of an offender will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices), then not only do we have an extra *pro tanto* reason to do so; in fact, in many cases we should do so, all things considered, because the values that would be compromised by mitigation of sentence do not generally (i.e. always or nearly always) trump the value of the improvement of the worst off.

However, I have *reservations* about the empirical claim of the argument. Some may be led to accept the empirical claim through the following simple reasoning: if we mitigate the punishment of a socially deprived offender, then we improve the well-being of him, who is a worst off member of society and whose disadvantage is unlikely to be justified by his responsible choices. (In the rest of this chapter I assume that for most worst off members of society, their disadvantage is not justified by their responsible choices.[258]) Accordingly, mitigating the punishment of socially deprived offenders will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices). In contrast, if we mitigate

[258] This claim is accepted by many writers on distributive justice. For example, see Barry, "Chance, Choice, and Justice", 149.

the punishment of a non-socially deprived offender, we do not benefit any worst off members of society (we only benefit a well-off person). Accordingly, the empirical claim is true. But, as I will show below, this simple reasoning is flawed at several points. Reducing the punishment of a socially deprived offender *sometimes will not* improve the overall well-being of the worst off members of society; and while it is less obvious, reducing the punishment of a non-socially deprived offender *may in many cases* improve the overall well-being of the worst off members of society. Therefore, while I do not have a knockdown objection against the empirical claim of the argument, *without further empirical data I do not believe we are entitled to accept it.*

But while, as just mentioned, I do not think that the improvement of the worst off argument is conclusively established, I believe it is worth discussing for the following two reasons. First, with a few notable exceptions,[259] many writers on punishment hold that punishment and distributive justice belong to distinct spheres, so even if we can benefit the worst off members of society by mitigating the punishment of some

[259] Ristroph and Hudson are sympathetic to my position but they did not defend this position in detail: see Ristroph, "Desert, Democracy, and Sentencing Reform", 1342, 1347; Hudson, *Penal Policy and Social Justice,* 15-16. Brown believes that punishment should be informed by cost-benefit analysis and "distributional" concern is said to be one of the relevant considerations. While he did not clarify what are the relevant distributional concerns it seems plausible to believe that our positions would share important similarities. See Brown, "Cost-benefit Analysis in the Criminal Law", 337, 353.

Interestingly, recently more arguments have been made for the claim that private law should be arranged in a way to further distributive justice. See Kordana and Tabachnick, "Rawls and Contract Law"; Keren-Paz, *Tort, Egalitarianism, and Distributive Justice.*

offenders there is no reason to mitigate their punishment, or that generally we should not do so all things considered.[260] In none of the major Western jurisdictions do judges mitigate the punishment of offenders on the ground that doing so will benefit the worst off members of society.[261] I believe that the improvement of the worst off argument is right in challenging that. Second, although I am not sure whether the proportion of socially deprived offenders who can rely on the idea of improvement of the worst off as a ground for mitigation of punishment is *greater* than that for those who are non-socially deprived offenders, I believe (and I will offer some justifications for this belief in section 4) that this idea can justify mitigation of sentence for many, or a significant proportion of, socially deprived offenders. So even if the improvement of the worst off argument cannot establish its empirical claim, what it has established may still have profound influence on sentencing of socially deprived offenders in the present distributively unjust society.

---

[260] See, e.g., Von Hirsch, *Censure and Sanctions*, 97-98, 108; Brodeur and Roberts, "Taking Justice Seriously", 91.

[261] The closest is the Canadian jurisprudence on the *Canadian Criminal Code* s 718.2(e), which requires that "all available sanctions other than imprisonment that are reasonable in the circumstances should be considered for all offenders, with particular attention to the circumstances of Aboriginal offenders". However, there are important differences between my position and the Canadian jurisprudence.

According to the Canadian jurisprudence, in order for an offender to receive a lenient treatment under the provision, he needs to show one of the following two things: first, his disadvantage is connected to the crime; or second, his cultural background makes a restorative justice sentence a more acceptable one to his community and thus more effective. (See, e.g. *Ipeelee* and *Gladue*.) In contrast to the first limb, according to the improvement of the worst off argument, whether we mitigate the punishment of an offender depends on the *effect* of mitigation but not on what contributed to the crime. In contrast to the second limb, while the improvement of the worst off argument looks at the effect of the sentence, the focus is not on its acceptability to the offender's community.

Moreover, one of the rationales for the Canadian provision is to reduce the overrepresentation of a certain group of people (aboriginals) in the prisons, but that is not one of the rationales of the improvement of the worst off argument. See section 3 of this chapter below.

The structure of this chapter is as follows. In section 2, I will defend the moral claim of the argument. In section 2.1, I will outline a conception of distributive justice, which can be called *responsibility-catering well-being prioritarianism.* In section 2.2, I will defend this conception of distributive justice. In section 2.3, I will explain that under this conception of distributive justice, if the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices) will be improved by mitigating the punishment of an offender, then distributive justice will be promoted to a greater extent by mitigating his punishment than mitigating the punishment of offenders the reduction of punishment of whom will not lead to that consequence. In section 2.4, I argue that distributive justice provides reasons for action for sentencing institutions such that if by mitigating the punishment of an offender we will promote distributive justice to a greater extent than mitigating the punishment of another offender, then there is a reason to mitigate the former offender's punishment that does not apply to the latter offender. After defending the moral claim of the improvement of the worst off argument, I will compare the improvement of the worst off argument with four related arguments in section 3.

In section 4, I will argue that not only do we have a *pro tanto* reason to mitigate an offender’s punishment if by doing so we will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices); in fact, in many cases we should do so, *all things considered,* because the values that would be compromised by doing so do not enjoy a general priority over the value of improving the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices).

In section 5, I will discuss the empirical claim of the argument. In section 6, I will examine whether socially deprived offenders will, in fact, be made worse off, all things considered, if distributive justice is taken into account in sentencing, as compared to a sentencing system that does not take into account distributive justice. In section 7, I will explore to what extent my main claims concerning the improvement of the worst off argument are robust in relation to different theories of distributive justice. In section 8, I will discuss whether my main claims are consistent with retributivism and crime reduction theories of punishment.

2. The moral claim of the argument

2.1. The conception of distributive justice that I hold: responsibility-catering

well-being prioritarianism as general distributive justice

In this section (2.1) I will lay out the conception of distributive justice that I hold without defending it. I will defend it in the next section (2.2).

First, let us distinguish between *general* goods and *local* goods, and accordingly, between general distributive justice (i.e. justice about the distribution of general goods) and local distributive justice (i.e. justice about the distribution of local goods). By "local goods" I mean goods that are constituted by the norms or modes of distribution and accordingly the social meanings of such.[262] Examples of local goods include "academic prizes" and "state punishment". By "general goods" I refer to goods like "resources", "primary goods", or "well-being". To illustrate the difference between the two kinds of goods, whether a good counts as an "academic prize" depends on the norms and the mode of distribution. If a student won a lottery from a company, then he received more utility, but he has not received more academic prizes. In contrast, whether something amounts to a general good like "utility" or "primary good" is largely independent from the question of how the good is acquired.

[262] Walzer, *Spheres of Justice*. See also Hild and Voorhoeve, "Equality of Opportunity and Opportunity Dominance", 120.

Many writers on distributive justice, like Bentham,[263] Rawls,[264] and Dworkin,[265] start by assuming that there is *an* important form of distributive justice that is general, i.e. there are valid principles about the distribution of general goods.[266] They disagree on what general distributive justice requires, i.e. what exactly is the *currency* of general distributive justice (primary goods, resources, welfare, or something else like capabilities?) and what exactly is the *principle* of general distributive justice (egalitarian, prioritarian, maximizing, or something else?).[267] My starting point on distributive justice is within that tradition (in contrast to, for example, a Walzerian theory that I will discuss in section 2.2 of this chapter below): I believe that there is *a* valid form of distributive justice that is concerned with the distribution of general goods. *When I refer to "distributive justice" without further elaboration in the rest of this thesis I should be understood as referring to "general distributive justice".*

(Let me stress at the outset that while I believe *general distributive justice is an important form of distributive justice*, I do *not* believe that it is the *only* valid form of distributive justice, nor do I believe that distributive justice is the only value relevant

[263] Bentham, *An Introduction to the Principles of Morals and Legislation.*
[264] Rawls, *A Theory of Justice: revised edition.*
[265] Dworkin, *Sovereign Virtue*, 1, 174-180, 401-406.
[266] But c.f. Cohen, *On the Currency of Egalitarian Justice*, 63, 113.
[267] For the distinction between *currency* of distributive justice and *principles* of distributive justice, see Arneson, "Welfare should be the Currency of Justice".

to distribution.[268] I will say more about these in section 2.2 of this chapter below.)

What then, should be the *currency* of general distributive justice, i.e. what is to be distributed? I believe that general distributive justice is concerned with the distribution of *well-being*. I want to make two clarifications on the concept of well-being I use. First, in the literature, there are three main ways to understand the concept of well-being: pleasure, desire-satisfaction, and objective account.[269] I adopt an objective account of well-being. Second, I believe we should adopt a pluralist account of objective well-being (i.e. we should adopt an objective *list* account). According to a pluralist account of well-being, a person's well-being is decided by how he fares in different *irreducible* dimensions, like friendship, achievement, love, pleasure, etc.[270] The different dimensions of well-being are irreducible as they concern incommensurable goods. Good A and good B are incommensurable if for a particular amount of good A and for a particular amount of good B, we do not have a rational basis to choose between them and a slight improvement in good A would not change our rational judgment.[271] If a person has better friendships, while the other

[268] So I do not accept, as Rawls claims, that "justice is the first virtue of social institutions". (Rawls, *A Theory of Justice: revised edition*, 3.) I think justice is an important value, but just one important value amongst others. See Cohen, "The Currency of Egalitarian Justice".
[269] For a general survey, see Griffin, *Well-being*.
[270] For an example of an objective list account of well-being, see Finnis, *Natural Law and Natural Rights.*
[271] Raz, *The Morality of Freedom*, chapter 13.

has better achievements, *sometimes* it is impossible to say who has a better life all things considered. That said, it is very important to observe that to hold that different dimensions of well-being are incommensurable does *not* mean that we can *never* make interpersonal *overall* well-being comparisons. Think about two persons. The first person has *vastly* more achievements and only marginally less health. Even if achievement and health are incommensurable, we may believe that overall speaking the first person's life is better.[272] Nor does the claim that different dimensions of well-being are incommensurable entail that we cannot improve the overall well-being of a person who suffers from a deficient of goods *in one dimension* by giving him *more goods in a different dimension*. To illustrate, suppose that a music-lover was made disabled by a car accident. Now, if we give him more money to buy CDs, he will not be restored to the status quo since the good of health and the good of aesthetic pleasure are incommensurable. But clearly his overall well-being will be improved by giving him money to buy CDs. (In this paragraph I am only concerned with arguing that the disabled music lover's *well-being can be improved* by giving him money. Whether distributive justice provides a reason to do so depends on whether there is a distributive justice reason to improve his well-being.)

[272] See, e.g. Hsieh, "Incommensurable Values", section 3.1.

Assuming that general distributive justice is concerned with well-being, what is the proper *principle* of distribution, i.e. what is a just distribution of well-being? I believe that responsibility-catering prioritarianism, offered by Arneson, is the answer: improving the well-being of a worse off member (if his disadvantage is not justified by his responsible choices) takes priority over improving the well-being of a better off member.[273] I will not give a particular account of how responsible choices can justify inequality here. That question is of course an important one though I do not think it is very crucial for my discussion in this chapter because I believe, no matter what exactly is the moral significance of responsible choices, generally speaking the disadvantage of the worst off members of our society cannot be justified by their responsible choices.[274]

How much priority does *distributive justice* require us to give to the worse off? Some prioritarians, like Arneson, believe that the priority is *not absolute* and that distributive justice only requires us to add a *weight* to the benefit on the worse off in our calculus.[275] To illustrate, suppose that we have to choose between benefiting a

[273] In general, my discussion on distributive justice in this chapter is greatly indebted to Arneson's work. Even the label, "responsibility-catering prioritarianism", is taken from his "Luck Egalitarianism and Prioritarianism". But there is one major difference between our views, concerning whether the priority given to the worse off is absolute insofar as justice is concerned. I will discuss this difference in the main text very soon.

[274] Barry, "Chance, Choice, and Justice", 149.

[275] Arneson, "Two Cheers for Capabilities", n15.

well off member of society greatly and benefiting a worse off member of society by a smaller amount (and assume that both persons are not responsible for their situations). General distributive justice, according to Arneson, will favour the former policy *if the improvement to the well-off person is great enough*.[276] On this particular point, I disagree with Arneson (whose theory of general distributive justice I largely adopt) and I am in agreement with Rawls: insofar as distributive *justice* is concerned, the priority is absolute: that is, distributive justice *always* requires giving the good to the worse off person.[277] In other words, I adopt a *maximin* version of prioritarianism as a matter of general distributive justice.

If what I said above is correct, then an important form of distributive justice (i.e. general distributive justice) requires us to act so as to maximize the amount of well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices).

2.2. Why responsibility-catering well-being prioritarianism?

In this section, I wish to defend my view on distributive justice, which I outlined in

[276] See Arneson, "Two Cheers for Capabilities"; see also Parfit, "Another Defence of the Priority View".
[277] For Rawls's difference principle, see Rawls, *A Theory of Justice: revised edition.*

the last section. My view has four main ideas: (1) a valid form of distributive justice is concerned with the distribution of general goods; in other words, general distributive justice is a valid form of justice; (2) the relevant kind of general goods to be distributed is objective well-being; (3) the relevant principle of distribution is responsibility-catering; and (4) the relevant principle of distribution is prioritarian; in addition, *insofar as* general distributive *justice* is concerned, the priority is an absolute one. Let me defend these four ideas in sections 2.2.1 to sections 2.2.4 respectively.

2.2.1. Is there any valid form of justice that is concerned with the distribution of general goods?

Many people, including myself, believe that there is a form of justice that is concerned with the distribution of general goods; the main task is to find out the correct currency and principle of general distributive justice. However, there are three notable views that disagree with that broad approach to the subject: libertarianism, Elizabeth Anderson's democratic equality theory, and a purely localist conception of distributive justice (which I take from the work of Michael Walzer). I will now explain these three views and provide reasons for rejecting them.

According to libertarianism, justice does not require us to achieve a particular distribution of general goods. Rather, whatever results from voluntary transactions after a just acquisition of initial resources is just.[278] I believe that libertarianism is mistaken but I cannot offer anything more than a very brief defence here. My reasons for rejecting libertarianism are given by many writers in the egalitarian tradition: voluntary transactions can always produce disadvantages that are not traceable to one's responsible choices, and in such cases the disadvantage seems to be morally arbitrary and it is hard to see why the worst off members of society have no cause, based on justice, to complain. In some cases there may be reasons, based on efficiency and freedom, for the state not to interfere with the result of voluntary transactions but it does not show that such result must be just.[279]

According to Anderson's democratic equality theory, the fundamental requirement of social justice is that everyone should stand in an equal social and political relationship. Contrary to responsibility-catering well-being prioritarianism and many other theories like utilitarianism and resource-egalitarianism, according to Anderson, there is no *intrinsic* requirement as to how general goods like well-being, primary goods, or resources should be distributed. Rather, we evaluate a pattern of distribution of

[278] Nozick, *Anarchy, State, and Utopia*.
[279] For persuasive critiques of Nozick's position, see Wolff, *Robert Nozick*; Cohen, *Self-Ownership, Freedom, and Equality.*

general goods based on whether it helps to achieve *equal social and political status* amongst citizens.[280]

Sometimes great inequality in some kinds of general goods will threaten equal social and political relationship. For example, in the real world economic power can easily be translated into social and political power. So in the real world, democratic equality would call for reducing inequality in economic goods, not because inequality in economic goods is bad in itself but rather because elimination of inequality in economic goods is a necessary means to reduce inequality in social relationships. However, in a world where inequality in general goods would not result in unequal political power or social status, there would be *no* cause for redistribution according to the democratic equality theory.

In reply, I agree with Anderson that there is a value in achieving equality in social and political relationships. But I also agree with Arneson that it is wrong to believe, as Anderson does, that there is *no* intrinsic requirement as to how general goods should be distributed.[281] That is, I believe it is not true that we should not care about the distribution of general goods if it does not affect whether people can stand in an equal

[280] Anderson, "What is the Point of Equality?". For a theory along a similar line, see Scheffler, "What is Egalitarianism?".
[281] Arneson, "Luck Egalitarianism and Prioritarianism".

social and political relationship. Suppose that two persons are in pain and we are deciding who should have a pain killer. The pain is serious but *not severe enough to prevent them from participating as equals in social and political relationships*. One of them, however, has suffered pain throughout his life through sheer bad luck while the other has a happy life through sheer good luck. I believe that there is a reason to give the pain killer to the first person but this result contradicts democratic equality, since the distribution of the pain killer would not affect whether the two persons stand as equals in society.[282]

According to a purely localist conception of justice (which I take from the work of Walzer),[283] most of the theories of distributive justice offered in the literature (Rawls' difference principle, responsibility-catering well-being prioritarianism, utilitarianism, equality of resources, etc.) are *too general*: they are principles concerning the distribution of general goods like well-being, resources, or primary goods. However, according to the purely localist conception, there is *no* valid principle governing the distribution of goods stated at that high level of generality – *all* valid principles of distribution are local, concerned with the distribution of local goods and burdens like education opportunities, academic prizes, state punishment, earnings in the workplace,

[282] Arneson, "Luck Egalitarianism and Prioritarianism", 347. See also Kymlicka, *Contemporary Political Philosophy: an Introduction: 2nd ed*, 199.
[283] Walzer, *Spheres of Justice*, 10.

household tasks, social security payment, and votes. The principle for distributing a kind of local goods is derived from the social meaning of that kind of local goods. For example, that academic prizes should be distributed according to academic merit is because of the social meaning of academic prizes.

One implication of the purely localist conception is that according to it, in contrast to a general conception of distributive justice, there is *no* reason to give a person who has received less than his proper share of *one kind* of local goods a greater share of *a different kind* of local goods to compensate. For example, suppose a poor person who had received less than his proper share of payments in the workplace entered an art competition with a monetary prize. According to most of the general conceptions of distributive justice, there will be a *defeasible* reason based on justice to give him the prize since the resulting distribution of general goods (resources/well-being/primary goods) may be closer to the ideal distribution (assuming that his competitors are rich and happy and assuming that *there is no alternative method* to give the poor person a greater share of resources/well-being/primary goods besides giving him the prize in the art competition). However, according to the purely localist conception, there is *no reason at all* to do so: doing so would not make justice any better (since the poor person would not thereby gain a fairer share of the first kind of local goods, namely, payment in the workplace); in fact, it would make justice *worse* (since he would get more than his fair share of the second kind of local goods, namely, art prizes).

In reply, I believe the purely localist conception of distributive justice is wrong. Let me begin by highlighting what seems to be the gem of insight in the purely localist conception: I agree that there *are* morally valid local distributive principles that are *distinct* from the general principle of distributive justice, i.e. there are valid principles governing the distribution of local goods *that are not derived* from the principles governing the distribution of general goods and may conflict with them. For example, I think that there is a reason for prizes to be distributed according to merit even if that does not improve the well-being of the worst off members of society. However, it seems to be mistaken to *deny*, as the purely localist conception wishes to, that there exist, *in additional to* local principles of justice, general principles of justice.[284] I want to make three arguments to support the claim that there is a form of distributive justice that is genuinely general.

The first argument is based on the Rawlsian idea that a form of justice is concerned with justifiability of the social arrangement to rational agents behind the veil of

[284] My position is similar to the one suggested in Holtug and Lippert-Rasmussen, "An Introduction to Contemporary Egalitarianism", 30. For another response to this Walzerian idea, see Arneson, "Welfare Should be the Currency of Justice".

ignorance.[285] A rational agent, in deciding whether to consent to a social arrangement in the original position, must look at the distribution of general goods like well-being in that social arrangement *since it is what matters most to them* once the veil is lifted. For example, if a rational agent behind the veil of ignorance knows that some people will not get their proper share of payment in the workplace and there is no way to compensate them other than to give them extra prizes in competitions, then he would opt for a scheme that would give those people extra prizes in competitions rather than a scheme that leaves their disadvantage uncompensated.[286]

The second argument essentially raises the same point but without appealing to the concept of the original position. According to this argument, justice requires *equal concern* of individuals.[287] In caring about an individual as a person, not simply as a worker or an artist, we do not only care about his performance in one dimension: we care about how he fares in his life *overall.* As such, it is insufficient for the state to ensure that local goods are distributed according to the local norms. In addition, there is a reason for the state to care about whether the distribution of general goods like well-being is proper.

[285] Rawls, *A Theory of Justice: revised edition.*
[286] See also Hudson, *Penal Policy and Social Justice*, 16.
[287] Dworkin, *Sovereign Virtue*, 1.

The third argument starts with the belief, which is held by many, that some redistributive policies are justified. It is hard to see how we can justify redistributive policies without any general principles of justice. Of course, *once* we accept that a certain amount of money is "welfare budget" then we can use local principles concerning the social meaning of "welfare goods" to justify distributing that amount of money to the poor. However, it is impossible to justify treating a particular amount of tax income of the state as "welfare budget" without a general principle of justice. This is because tax income is just money – it does not come with a particular social meaning so it is hard to see how any local principles can justify treating a particular amount of tax income as "welfare budget".[288]

Based on the three arguments above, I believe that there are valid principles of justice governing the distribution of general goods *in addition to* valid principles governing the distribution of local goods. I believe that also does justice to our moral phenomenology. Recall the case above about whether to give the poor person extra prizes in the competition. *If* the purely localist conception is correct, then that case is an *easy* case: justice *unequivocally* demands not giving him extra prizes. But I believe that many people take that to be a case involving some kind of *conflict:* there is

[288] For a similar point, see Barry, "Intimations of Justice", 809-810.

something to regret no matter which course of action we take. I feel that it is unfair, in some sense, to give the poor person extra prizes, *but also unfair,* in some sense, to not give him extra prizes. Only if there exist general principles of distributive justice can we explain the sense of unease there.[289]

Let me stress that all I have tried to show, in my reply to the purely localist conception above, is that general distributive justice is *a* valid form of justice. I have *not* denied (and in fact I have accepted) that there are, in addition to general principles of distributive justice, local principles of distributive justice, which may conflict with general principles in a particular case (the case above about giving art prizes to the poor person who had not received his proper share of payment in the workplace is precisely a case where the principle concerning the distribution of general goods conflicts with the principle concerning the distribution of prizes in art competitions). Moreover, I have not yet discussed the relative weight and priority between local and general principles. I will argue in section 4 of this chapter that there is no reason to believe that the local principle concerning the distribution of punishment should always or nearly always trump the general principle concerning the distribution of general goods (like well-being) but this point has not been argued for yet.

[289] Affirmative action in education may be a case for a similar point. See Arneson, "Against Complex Equality", 104.

2.2.2. Why objective well-being as the currency of general distributive justice?

Now let me explain why the currency of general distributive justice is objective well-being rather than other accounts of well-being, like satisfaction of preferences, or general goods other than well-being, like primary goods. The reason is that I agree with Arneson that general distributive justice should care about the distribution of things that *ultimately matter to people prudentially*, and what ultimately matters to a person prudentially is his objective well-being. What we should care about, prudentially, is not simply how much pleasure we have in our lives or how many of our preferences have been fulfilled, but rather how our lives have been objectively. A life full of pleasure in a pleasure machine is not a good one, nor is a life full of satisfaction of valueless desires. Similarly, it seems unmotivated to care about resources or primary goods if resources or primary goods will not improve the agent's well-being.[290]

[290] Arneson, "Welfare should be the Currency of Justice".

2.2.3. Why responsibility-catering?

Let me explain why I believe a proper theory of general distributive justice should be responsibility-catering (i.e. it asks us to compensate a disadvantaged person *only if* the extent of his disadvantage cannot be justified by his responsible choices). The main reason to accept this constraint is simply that it is intuitive: it does seem plausible that if a person deliberately gambles with his well-being, or sacrifices himself, then he is entitled to less as a matter of distributive justice.[291]

For the sake of simplicity, when I refer to "worst off members of society" in the rest of this chapter, unless the context clearly suggests otherwise I mean "worst off members of society (whose disadvantage is not justified by their responsible choices)".

2.2.4. Why prioritarianism?

Now let me explain why the principle of general distributive justice should be prioritarian, rather than maximizing, egalitarian, or sufficientarian. The idea that

---

[291] See, e.g. Cohen, "The Currency of Egalitarian Justice".

disadvantage, at least if unchosen, ought to be compensated is a considered conviction that many people, including myself, share.[292] This idea also can be defended by appealing to Rawls's idea of the original position.[293] Prioritarianism is precisely the principle that calls for removal of disadvantage and therefore is attractive. A maximizing principle of distribution would leave unchosen disadvantage uncompensated – it may even call for increasing the disadvantage -- and it therefore fails to treat everyone with equal concern. Of course, that does *not* mean that we should always maximize the well-being of the worst off, *all things considered,* since distributive justice may give way to other values, like utility, sometimes. But in such a case the honest thing to say is that distributive justice is given up for other values rather than that justice requires sacrificing the worst off.

Egalitarianism, like prioritarianism, compensates disadvantage. However, egalitarianism, unlike prioritarianism, assigns *negative* value to an improvement of the better off and accordingly it is vulnerable to the levelling down objection offered by Derek Parfit.[294]. For example, harming the position of Bill Gates would promote distributive justice according the egalitarianism but it will harm distributive justice

[292] Arneson, "Luck Egalitarianism and Prioritarianism".
[293] Rawls, *A Theory of Justice: revised edition.*
[294] For the levelling down objection, see Parfit, "Equality or Priority". For a response, see Temkin, "Equality, Priority and the Levelling Down Objection".

(albeit only insignificantly) according to prioritarianism. I believe that provides a good reason to favour prioritarianism over egalitarianism.

Sufficientarianism is different from prioritarianism because an improvement of the position of a person above a certain sufficiency threshold, according to sufficientarianism, does not promote distributive justice at all.[295] For example, improvement of the position of Bill Gates would not promote distributive justice according to sufficientarianism, but it will promote distributive justice (albeit only insignificantly) according to prioritarianism. The main difficulty with sufficientarianism, which is a well-known one and explains why I favour prioritarianism over it, is that it is very hard to specify the line of sufficiency in a non-arbitrary manner.[296]

I do not simply believe that general distributive justice requires *some* priority to be given to the worst off; I believe that, *insofar as general distributive justice is concerned*, the priority is *absolute.* To illustrate my point, suppose we are choosing between benefitting a better off by a large amount, and benefiting a worse off by a smaller amount. I believe that distributive justice *always* favours the latter policy no

[295] See, e.g. Frankfurt, "Equality as a Moral Ideal".
[296] Casal, "Sufficiency is not Good Enough"; Arneson, "Distributive Justice and Basic Capability Equality: 'Good Enough' Is Not Good Enough".

matter how much benefit we can confer to the better off. What are the justifications for my belief? This has to do with the rationale for prioritarianism I just offered: unchosen disadvantage is morally arbitrary and hence unjust. The worse off always has a complaint if we choose the former policy no matter how great is the benefit to the better off, as his unchosen disadvantage could otherwise be reduced. So I believe that *from the point of view of justice* we should always favour the latter policy. Again, that does *not* mean we should always choose the latter policy *all things considered* – but, as I said above, the honest thing to say in a case where the former policy is chosen is that justice is given up for other values like utility rather than that justice calls for sacrificing the worst off members of society. That seems to be the right thing to say: if we choose the former policy I believe that we have a duty, at least, to apologize to the worse off, which is hard to explain if no injustice has been done against him.[297]

That concludes my brief defence of my view on distributive justice. I understand that the issue of what distributive justice requires is one of the most controversial topics in political philosophy and some people will not be persuaded by everything I just said in sections 2.1 and 2.2. So in section 7 of this chapter, I will explore to what extent

[297] For similar ideas on residual duty, see Gardner, "In Defence of Defences", 82; Simmons, *Moral Principles and Political Obligations*, 27-28.

my main claims concerning the improvement of the worst off argument are robust in relation to different views on distributive justice.

2.3. If by mitigating an offender's punishment the well-being of the worst off members of society will be improved, then distributive justice will be promoted to a greater extent by mitigating his punishment than mitigating the punishment of offenders the reduction of punishment of whom will not produce that consequence

If responsibility-catering well-being prioritarianism, a theory that gives priority to the improvement of the well-being of the worst off members of society, is the correct theory of general distributive justice, then it follows that general distributive justice will be promoted to a greater extent by a measure that will produce that consequence as compared to a measure that will not. So if the well-being of the worst off members of society will be improved by mitigating the punishment of an offender, then general distributive justice will be promoted *to a greater extent* by mitigating his punishment as compared to mitigating the punishment of offenders the reduction of punishment of whom will not lead to that consequence.

Let me make one point in clarification. General distributive justice, under prioritarianism, *can* be promoted in ways *other than* improving the worst off. This is because under prioritarianism, unlike egalitarianism (which assigns negative value to the improvement of the well off) and sufficientarianism (which assigns zero value to the improvement of the well off if they are above the line of sufficiency), distributive justice will also be promoted if the position of the *well off, rather than the worst off,* is improved. So it is *not* true that, under prioritarianism, the *only* way to promote general distributive justice through mitigation of punishment is by mitigating the punishment of offenders the reduction of punishment of whom will lead to an improvement of the well-being of the worst off.

That said, under maximin prioritarianism, *the extent* of general distributive justice promoted by a measure that will only improve the position of the well off *is always less* than the extent of general distributive justice promoted by a measure that will improve the well-being of the worst off. Therefore, the fact that we *can* promote general distributive justice by mitigating the punishment of offenders the reduction of punishment of whom will *not* lead to an improvement of the well-being of the worst off does not affect the point that we *always* promote general distributive justice *to a greater extent* by mitigating the punishment of offenders the reduction of punishment of whom will lead to an improvement of the well-being of the worst off.

2.4. If by mitigating an offender's punishment we will promote general distributive justice to a greater extent than mitigating the punishment of another offender, then there is a reason to mitigate the former's punishment that does not apply to the latter

If by mitigating an offender's punishment we will promote general distributive justice to a greater extent, then there seems to be an extra reason to mitigate his punishment as compared to other offenders as general distributive justice is a value and it is plausible to claim that general distributive justice is a value which, unlike some other values say promotion of virtue, is a legitimate concern of the state.

Here my opponent may reply that general distributive justice, while being a value that should be taken into account by *the state*, is not a value that should be taken into account *in sentencing* at all;[298] hence, even if general distributive justice will be promoted to a greater extent by mitigating an offender's punishment there is no extra reason to mitigate his punishment; and accordingly there is no reason to punish him less than an otherwise similar offender the reduction of punishment of whom can only

[298] Roberts, "The Sentencing of Aboriginal Offenders in Canada: A Rejoinder".

promote distributive justice to a lesser extent. But what are the arguments for claiming that general distributive justice, while providing reasons for action for the *state*, does not provide reasons for action *for sentencing institutions, a particular set of state institutions*?

I can think of four such arguments. The first argument is based on the idea that general distributive justice does not form part of the justification of punishment. The second argument appeals to the existence of other institutions that can be more efficient in promoting general distributive justice. The third argument appeals to political feasibility. The fourth argument appeals to the original intent in creating the sentencing institutions. I will argue that none of these four arguments shows that general distributive justice does not provide reasons for action for sentencing institutions.

2.4.1. Not part of justification of punishment objection

The first argument for the claim that we should not use punitive institutions to promote general distributive justice is that general distributive justice is not part of the justification of punishment and accordingly does not provide reasons for action for punitive institutions. One presupposition of this argument is that general distributive justice is not a justification of punishment, and writers like Murphy will dispute that presupposition.[299] I, however, do not want to question that presupposition here. Instead I want to challenge another point of this argument, the claim that if a value is not part of the justification of an institution, then there is no reason for that institution to promote that value.

I think Hart is right in claiming that values that do not form part of the justification of an institution can matter as to how the institution should act.[300] To illustrate this point, consider the constraint against killing innocent enemy civilians in war. The justification for going to war is not to prevent the death of civilians of the enemy state but rather, say, self-defence. The fact that we should sometimes refrain from killing enemy civilians in war, even if by killing them we can increase our chance of achieving the purpose of the war, provides an example in which an institution's actions should be affected by a value that is not amongst one of its purposes.

2.4.2 The other institutions objection

[299] Murphy, "Marxism and Retribution"; Sadurski, "Distributive Justice and the Theory of Punishment".
[300] Hart, "Prolegomena to the Principles of Punishment".

According to the second objection, the *other institutions objection*, there are institutions besides the sentencing institutions, like the welfare institutions, that are better equipped to improve the well-being of the worst off members of society.[301] As such, general distributive justice only provides a reason for improving our welfare institutions, say through a comprehensive welfare reform, but not a reason to promote general distributive justice in a haphazard manner through the sentencing system.

The problem of this objection is that *the alleged alternative means to promote general distributive justice is generally unavailable in the real world*. It is certainly unrealistic to think that, given how many resources our societies are willing to spend on welfare policies in the foreseeable future, our welfare policies can fully compensate the worst off members of society. I concede that in *an ideal world*, where the worst off members of our society will be compensated fully through institutions outside the sentencing system, general distributive justice would not provide a reason to alter the sentencing system. But that does not mean that general distributive justice does not provide a reason for us to alter our sentencing system i*n the real and non-ideal world* where the worst off members of society cannot be compensated otherwise.

301 Morse, “Deprivation and Desert”, 115. A similar point against a slightly different proposal can be found in Parent, “The Whole Life View of Criminal Desert”, 353.

An analogy with the law granting limited rights to homosexuals in a civil partnership can help to illustrate my point. For many people, homosexual couples in civil partnership should enjoy full legal rights as heterosexual couples in marriage. But that kind of law is not likely to be enacted in many societies in the real world; and for those societies, the fact that a law granting limited rights to homosexuals is not the ideal solution does not mean that there is no reason to try to get the law enacted, if the ideal solution is unattainable. Analogously, the fact that sentencing reform is not the best way to improve the well-being of the worst off members of our society does not show that there is no reason for that based on general distributive justice if the ideal solution to eliminate general distributive injustice, that is, a comprehensive welfare reform, which provides for the worst off members of our society outside the sentencing system, is not available in the reasonably foreseeable future.[302]

2.4.3. Political feasibility

A value only provides a reason to attempt an act if the act is possible. Aesthetic value will be promoted if we can resurrect Beethoven. But there is no reason for us to try to do that because it is not a physically possible option for us.

---

[302] This point is essentially made by Alan Norrie in Norrie, "From Criminal Law to Legal Theory: The Mysterious Case of the Reasonable Glue Sniffer", 544 n29. For expression of similar sentiments, see Carter, "Of Fairness and Faulkner", 63.

This idea, stemming from the common thought that "ought implies can", may ground another objection against using sentencing reform to promote general distributive justice. It may be argued that while using sentencing reform to improve the well-being of the worst off members of society is not physically impossible,[303] it is not politically feasible, i.e. it is unlikely to be passed by the legislature.[304]

In reply, I accept that there is no reason to embark on the impossible. But why should we believe that using sentencing reform to improve the well-being of the worst off members of society is politically impossible? The political feasibility of such measure can, obviously, vary from society to society, and I cannot see why, as a general matter, we should be as pessimistic as my opponent. In fact, one example may help to illustrate that such a policy need not be politically impossible in most contemporary Western societies. The Canadian legislature has enacted a statute that requires judges to take into account "the circumstances of aboriginal offenders" in sentencing.[305] While there are important differences between the relevant Canadian provision and the improvement of the worst off argument,[306] I believe the Canadian experience

[303] Some people may argue that while it is physically possible to promote distributive justice through sentencing, the amount of distributive justice we can promote through sentencing is minimal. I will address this concern in section 4 of this chapter below.
[304] Von Hirsch, *Censure and Sanctions,* 108; von Hirsch and Ashworth, *Proportionate Sentencing,* 72.
[305] See *Canadian Criminal Code* s 718.2(e).
[306] See section 1 of this chapter above.

shows that legislatures are not always hostile to taking into account broader social considerations in sentencing.

But perhaps my opponent may raise another objection. Instead of focusing on *absolute* political feasibility, he may focus now on *relative* political feasibility. By that I mean that my opponent may focus on the relative political feasibility between the following two proposals:

(1) sentencing reform that promotes general distributive justice (my proposal).

(2) comprehensive welfare reform.

My opponent may argue that proposal (1) is not more politically feasible than proposal (2), and hence, there is no reason to go for proposal (1). Why does the fact that proposal (2) is more politically feasible than proposal (1) show that there is no reason to go for proposal (1)? My opponent may argue that this is because of the fact that proposal (1) is a *second best solution* to improve the well-being of the worst off members of society. Recall that according to me, the best way to improve the well-being of the worst off members of society is proposal (2). It is only because proposal (2) is unlikely to be adopted in the real world that we should go for proposal

(1) as the non-ideal solution.[307] According to my opponent, once we understand that proposal (2) is the ideal solution and proposal (1) is the non-ideal second best solution, we should accept that there is a reason to go for proposal (1) only if it is more politically feasible than proposal (2): there is no reason to go for a second best solution if the second best solution is as politically infeasible as the ideal solution.

In reply, I agree with my opponent that there is a reason to go for proposal (1) only if proposal (1) is more politically feasible than proposal (2). But I also believe that is actually the case for many societies. The main reason for my belief is that, whatever the *moral* costs of proposal (1) (it involves, for example, sacrificing a certain degree of proportionality, an issue to be discussed in section 4 of this chapter), it is much *less expensive, in material terms,* than a comprehensive welfare reform. (In fact, proposal (1) may actually enable the state to spend less money through reduction of prison population.[308]) This is important as I believe that many people object to a comprehensive welfare reform because of its material cost. To these people proposal (1) would be more acceptable than proposal (2). As such proposal (1) can be a more politically feasible means to improve the well-being of the worst off members of

[307] There are other ways of drawing the distinction between ideal and non-ideal theories. For a recent discussion, see Simmons, "Ideal and Non-ideal Theory".
[308] Though, as Ashworth correctly reminded us, the effect of reduction in prison population caused by sentencing reform should not be overestimated, given that sentencing is only one out of the many parts of the penal system. Ashworth, *Sentencing and Criminal Justice: fifth edition,* 22.

society and accordingly general distributive justice than proposal (2).

2.4.4. Original intent in creating sentencing institutions

The fourth objection against the claim that there is a reason to use sentencing to promote general distributive justice, which can be called the *original intent objection*, argues that since sentencing institutions were not *originally designed* to promote general distributive justice, general distributive justice is none of their business.[309]

The main problem with this objection is that it is unclear why the original intent in creating an institution matters in deciding what values should the institution care about. My opponent may raise two suggestions to explain the connection. First, if an institution is not designed to promote a value, then it may not be very efficient for the institution to promote that value, such that great cost is involved in using the institution to promote that value. Second, if an institution is not designed to promote a value, then other institutions are probably designed to promote that value so promotion of that value should be left as a business for other institutions.

[309] This idea was suggested to me by someone else, but I am afraid that I cannot trace the exact source at this moment. Morse, "Deprivation and Desert", 153-154 seems to hint at a similar idea.

In reply, I think both suggestions fail. Against the first suggestion one can raise two points. First, the fact that an institution was not originally designed to promote a value does not entail that it is inefficient for the institution to promote that value. That is certainly a contingent matter. Looking at our issue more specifically, my opponents need to explain why using mitigation of sentences to promote general distributive justice is inefficient, given that reducing sentences on certain offenders does not look like a very costly matter. (I will revisit this issue in section 4 below.) Second, even if sentencing is not an efficient means to promote general distributive justice, it only shows that *all things considered*, we should not use sentencing to promote general distributive justice. It does not show that there is *no reason* to do so.

The second suggestion is basically the same as the other institutions objection, and my reply has been stated above: the fact that there are institutions, other than the sentencing institution, that are better equipped to promote general distributive justice does not mean that there is no reason to use the sentencing institution to promote general distributive justice when those other institutions are not doing their work.

3. Comparing the improvement of the worst off argument with four other arguments

I hope, after going through the moral claim of the improvement of the worst off argument, my readers would by now have a good grasp of the argument. I wish now to clarify the improvement of the worst off argument by comparing it with four other non-exculpatory non-communicative arguments. The four arguments are a utilitarian argument for mitigation of sentence, an argument made by Murphy based on unfair advantage, an argument that appeals to the idea of whole life moral desert, and an argument based on correcting the overrepresentation of poor people in the prisons.

Let us start by comparing the improvement of the worst off argument with the utilitarian idea that whether we should reduce the punishment of some offenders depends on whether we will maximize aggregate utility in doing so. The improvement of the worst off argument is, *like the utilitarian idea, a consequentialist argument.* According to the improvement of the worst off argument, the reason why there is a mitigating factor that is more applicable to socially deprived offenders is because mitigating the punishment of a socially deprived offender is more likely to produce a particular *good consequence,* namely, an improvement of the well-being of the worst off, than mitigating the punishment of a non-socially deprived offender. However, the improvement of the worst off argument is also importantly different from a utilitarian basis of mitigation, since *the good consequences that form the focus of the two arguments are different:* a measure that increases aggregate utility need not improve the well-being of the worst off members of society, and vice versa.[310]

I believe, for two reasons, that the improvement of the worst off argument is more promising than the utilitarian idea in justifying a mitigating factor that has a greater applicability to socially deprived offenders. First, there seems to be absolutely no basis for thinking that mitigating the punishment of a socially deprived offender is more likely to result in a greater increase in aggregate utility than mitigating the punishment of a non-socially deprived offender (except cases where the punishment is a fine). If anything, a day outside prison is often more enjoyable for a rich person than a poor person, so reducing the same amount of punishment of the rich person may increase aggregate utility to a greater extent. In contrast, as I mentioned in section 1 of this chapter above, there is at least an intuitive line of thought (which I will ultimately doubt though in section 5 of this chapter below) for claiming that mitigating the punishment of a socially deprived offender is more likely to improve the well-being

[310] For the distinction between utilitarianism and consequentialism, see Sen and Williams, "Introduction".

of the worst off members of society than mitigating the punishment of a non-socially deprived offender.

Second, as I mentioned in section 2 of this chapter, improving the well-being of the worst off members of society is generally a much more important moral concern than an increase in aggregate utility. A measure that benefits the poorest is much more urgent than a measure that benefits the rich if the amount of benefit involved is the same. Accordingly, if the basis for mitigating an offender rests not simply on an increase in aggregate utility but rather on the improvement of the worst off, then there is a much stronger case to do that all things considered.

(The moral claim of the improvement of the worst off argument is also different from the view that whether we should mitigate the punishment of an offender depends on whether we maximize *weighted* utility in doing so. By weighted utility, I mean that a unit of utility counts more if it is on someone who is worse off. This weighted utilitarian view on mitigation shares similarities with the moral claim of the improvement of the worst off argument and may justify mitigation for similar offenders in practice (depending on the weight adopted by such view). However, they are still conceptually different, as I explained in section 2.1 of this chapter above. As long as the weight is not absolute, a measure that improves the well-being of the worst off need not maximize weighted utility and vice versa.)

Let us move on to the famous argument made by Murphy for claiming that socially deprived offenders and non-socially deprived offenders should be treated differently in punishment, an argument that actually provides the inspiration for the improvement of the worst off argument. Murphy starts by defending a theory about justification of punishment. According to Murphy, punishment is justified as *an aspect of distributive justice*: the right and reason to punish come from the following: a) non-socially deprived offenders gained unfair advantage through their crimes, b) distributive justice requires us to eliminate unfair advantage, and c) punishment is a means to do so. It then follows that socially deprived offenders should be punished less than non-socially deprived offenders: since socially deprived offenders do not, because of their past disadvantage, have a *net* unfair advantage over other citizens even after obtaining the benefit through their offences, there is no reason based on distributive justice to punish them, unlike non-socially deprived offenders.[311]

An objection to Murphy's argument, which I will call the *unclear benefit objection*,

[311] Murphy, "Marxism and Retribution".

has been raised by many. According to the unclear benefit objection, it is unclear what benefits has an offender, even a non-socially deprived one, gained by offending and that punishment is a proper way to eliminate the unfair advantage.[312] Two suggestions as to the nature of benefit obtained by offending are possible, and neither seems satisfactory. The first suggestion appeals to gain in welfare or material goods (like the proceeds of crime). But this suggestion clearly does not work as there are many crimes (e.g. criminal damage) that are punishable but need not produce any gain in welfare or material goods for the offender.[313]

The second suggestion, which is Murphy's own view, appeals to the offender's gain in freedom, through disobedience of the law.[314] According to this answer, everyone incurs a cost in freedom by exercising their restraint in obeying the legal rules. If an offender disobeyed, he gained something, namely, extra freedom, even if he did not gain any welfare or material goods.[315] One can attack the second suggestion by claiming that it is unclear that such freedom can really be a form of benefit or advantage.[316] Another route to attack this suggestion was offered by Victor Tadros. If the unfair advantage gained by the offender consists in freedom from self-restraint

[312] Odudu, "Retributive Justice in an Unjust Society", 419.
[313] Boonin, *The Problem of Punishment*, 121.
[314] Murphy, "Marxism and Retribution", 54.
[315] Bonnin, *The Problem of Punishment*, 121.
[316] For related problems see Boonin, *The Problem of Punishment*, 122-135; Odudu, "Retributive Justice in an Unjust Society".

and such gain in freedom is conceptually distinct from gain in welfare, then the moral principle that calls for elimination of unfair advantage would only ask us to reduce the offender's freedom and would not call for any reduction of the offender's welfare. But a measure that only limits the offender's freedom but which has no impact on the offender's welfare at all is not a paradigmatic case of punishment. Hence, it is unclear how we can justify punishment as a means to eliminate unfair gain in freedom.[317]

The unclear benefit objection is a strong objection against Murphy's argument. The improvement of the worst off argument, I believe, enjoys one advantage over Murphy's argument because it avoids the unclear benefit objection. Like Murphy's argument, the improvement of the worst off argument is based on general distributive justice[318] and it agrees with Murphy that general distributive justice provides reasons for action for punitive institutions. Unlike Murphy's argument, the improvement of the worst off argument does not need to commit to (although it also does not need to reject) the idea that distributive justice is part of the justification of punishment. Rather, the improvement of the worst off argument claims (plausibly, as I have argued in section 2.4.1 of this chapter above) that *even if* distributive justice is not part of the *justification* of punishment, it is still relevant to the *distribution* of punishment and

[317] Tadros, *The Ends of Harm*, 86-87.
[318] Though the particular conception of general distributive justice presupposed by the improvement of the worst off argument may be slightly different from that presupposed by Murphy's argument.

hence still provides reasons for action for punitive institutions. Because the unclear benefit objection is an objection against the claim that general distributive justice forms a *justification* of punishment but it is not an objection against the claim that general distributive justice should be taken into account in deciding the *distribution* of punishment, it is a good objection against Murphy's argument but not against the improvement of the worst off argument.

The improvement of the worst off argument is also different from an argument based on the idea of *whole life moral desert*. According to the idea of whole life moral desert, the state of affairs, where the total level of misery in each person's whole life corresponds to the total level of moral blameworthiness displayed in his whole life, is a good one. [319] Since socially deprived offenders have already suffered a lot in their lives, reducing their punishment will help to promote whole life moral desert, while that is not the case for mitigating the punishment of non-socially deprived offenders. I will discuss the pros and cons of this whole life moral desert argument in chapter 10 of this thesis. Here, I only wish to observe the difference between the two arguments. For the idea of whole life moral desert, how much happiness or suffering one should get depends on his *moral blameworthiness*. For general distributive justice, in contrast,

[319] Ezorsky, "The Ethics of Punishment", xxii-xxvii.

the relevant input is usually thought to be choices and efforts rather than moral blameworthiness. Many people believe that a lazy but otherwise saintly person is not entitled to more goods, *as a matter of general distributive justice,* than a hard working but morally average person.[320]

The improvement of the worst off argument is also different from the argument that socially deprived offenders should be punished less in order to correct the overrepresentation of socially deprived people in our prisons. Reducing the overrepresentation of aboriginal people in our prisons has been offered as one of the justifications for punishing them less in the famous Canadian case *R v Gladue.*[321] That argument is grounded on correcting unfair distribution of *punishment* based on population proportion and hence is different from the improvement of the worst off argument, which is grounded on correcting unfair distribution of *well-being.* Under the improvement of the worst off argument, if by punishing a socially deprived offender less we can improve the well-being of the worst off members of society, then there is a reason to do so even if it happens that socially deprived people are *under*represented in our prisons based on their population proportion.

[320] Hurka, "Desert: Holistic and Individualistic".
[321] For another argument based on correcting overrepresentation of poor people in our prisons, see Butler, "Class-based Remedies for the Poor", 225-226.

4. In many cases, we should reduce an offender's punishment, all things considered, if by doing so we will improve the well-being of the worst off members of society

Even if there is an extra *pro tanto* reason to punish an offender less for his crime if by doing so we will improve the well-being of the worst off members of society, generally (i.e. always or nearly always) we should not do so, all things considered, if the values that would be compromised by such mitigation of sentence generally outweigh the benefits. I want to suggest, however, that it is not the case by attacking six arguments against mitigation. The first argument claims that the amount of well-being of the worst off members of society that can be improved by mitigation of punishment is minimal. The second to the sixth arguments claim that the cost (moral or material) in mitigating punishment for improving the well-being of the worst off members of society is very large.

According to the first argument, we can at most bring about minimal improvement to the well-being of the worst off members of society through the sentencing system: it is frequently said that sentencing is not a good tool to achieve distributive justice.[322]

[322] Von Hirsch, *Censure and Sanctions*, 97-98, 108; Brodeur and Roberts, "Taking Justice Seriously", 91; Anand, "The Sentencing of Aboriginal Offenders, Continued Confusion and Persisting Problems:

However, why should we believe that? My opponent may argue it is because first, an offender's well-being will not be substantially affected by the severity of sentence, and second, most worst off members of society do not offend, so mitigation of sentence can only help a very small amount of worst off members of society.

Two points can be made in reply. First, in deciding how much improvement to the well being of the worst off we can bring about by mitigating punishment, we need to look at the impact of mitigation of punishment on people *other than* the offenders whose punishment is mitigated. Even if we do not look at broader social consequences caused by the mitigation of sentence,[323] at least the family of the offender may benefit from it in addition to the offender.[324]

Second, it is unclear that mitigation of sentence cannot substantially improve the well-being of the worst off members of society even if we focus exclusively on the impact on offenders whose punishment is mitigated. A lengthier term of prison or a

---

A Comment on the Decision in R. v. Gladue".

[323] See, e.g., Fagan, West, and Holland, "Reciprocal Effects of Crime and Incarceration in New York City Neighborhoods".

[324] Clear, "The Effects of High Imprisonment Rates on Communities"; Levy-Pounds, "From the Frying Pan into the Fire: How Poor Women of Color and Children are Affected by Sentencing Guidelines and Mandatory Minimums"; Cudjoe and Barringer, "More than Mere Ripples: The Interwoven Complexity of Female Incarceration and the African-American Family"; Wilderman, "Parental Imprisonment, the Prison Boom, and the Concentration of Childhood Disadvantage"; Murray and Farrington, "The Effects of Parental Imprisonment on Children"; Editors of the Yale Law Journal, "On Prisoners and Parenting: Preserving the Tie that Binds".

I am only claiming that mitigation of sentence can *sometimes* benefit the family of the offender and I do not deny sometimes it can have the contrary effect. See Wakefield and Wilderman, *Children of the Prison Boom*, chapter 4.

heavier fine significantly reduces freedom, which affects one's well-being.[325] Moreover, sometimes a change in the *form* of sanction may have substantial impact on the long term life of the offender. For example, there is some empirical evidence to suggest that an offender's long term economic prospects will be negatively affected by a custodial sentence as compared to a non-custodial sentence.[326] Concerning the point that most worst off members of our society do not offend, I wish to observe that in fact a lot of worst off members of our society are involved with the criminal justice system. A study done in 2003 by the US Bureau of Justice is instructive.[327] According to this study, it is estimated that one third of black people born in America in 2001 will be imprisoned at least once in their lives. Of course, race is not exactly co-extensive with one's amount of well-being and the extent to which worst off members of society commit crimes will vary from one society to the next. But it seems plausible to believe that, for most Western Societies, a significant proportion of their worst off members would be directly affected by criminal sentencing.

It is sometimes said that sentencing policy cannot eliminate the root of distributive

325 On the idea that money is a form of freedom, see Cohen, "Freedom and Money". It should be noted that subjective happiness is only one of the many dimensions of well-being. Hence Bronstein et. al.'s work, which shows that sentence length is irrelevant to the subjective happiness of the offender, does not show that sentence length is irrelevant to his well-being. See Bronstein et. al., "Happiness and Punishment".

326 Grogger, "The Effect of Arrests on the Employment and Earnings of Young Men"; Waldfogel, "The Effect of Criminal Conviction on Income and the Trust Reposes in the Workmen". For a survey of related studies, see Western, Kling, and Weiman, "The Labor Market Consequences of Incarceration". But for a contrary study, see Loeffler, "Does Imprisonment Alter the Life Course?".

327 US Bureau of Justice Statistics, *Prevalence of Imprisonment in the US population, 1974-2001.*

injustice.[328] That is true, but it does not follow that mitigation of sentence cannot substantially promote distributive justice. Sending out the full punishment in many cases can substantially worsen distributive justice. Mitigation of punishment in those cases can, therefore, promote distributive justice by not harming distributive justice further.[329]

The second argument claims that to mitigate an offender's punishment to improve the well-being of the worst off members of society would violate the principle of *proportionality* and proportionality is more important than improving the well-being of the worst off members of society. The principle of proportionality is, as I mentioned in chapter 1, the requirement that the severity of punishment imposed for a crime should be proportionate to the offender's blameworthiness for it. Since the question of whether mitigating an offender's punishment will improve the well-being of the worst off members of society is different from the question of how blameworthy he is for his crime, if we reduce an offender's punishment simply because by doing so we will improve the well-being of the worst off members of society we would violate the principle of proportionality.[330]

328 See, e.g., McCoy, "Sentencing (and) the Underclass", 592.
329 For a similar point, see Harris, "A Brief for De-escalating Criminal Sanctions", 216.
330 Von Hirsch and Ashworth, *Proportionate Sentencing,* 70-71.

In reply, I agree that using mitigation of punishment to improve the well-being of the worst off members of society would require some sentences to depart from proportionality. But while I accept that proportionality is a valid and important moral principle in punishment (see chapter 1 of my thesis), I do not agree that proportionality is generally more important than improving the well-being of the worst off members of society.[331]

There are three main attempts to justify proportionality and I believe none of them highlights a value that generally trumps the importance of improving the well-being of the worst off members of society. The first is based on expressive concerns: the severity of punishment expresses the degree of censure and the degree of censure should be proportionate – otherwise the state would be expressing a false message.[332] But even if we accept that proportionality can be justified by expressive considerations, I cannot see why expressive considerations are generally more important than the moral value of improving the well-being of the worst off members of society.

---

[331] I echo Michael Tonry's position on this point. See Tonry, "Proportionality, Parsimony, and Interchangeability of Punishments", 142. See also Roach and Rudins, "Broken Promises". It is unclear though how absolute defenders of the principle of proportionality take that principle to be. See von Hirsch, *Censure and Sanctions*, 47-56; von Hirsch, "Sentencing Reform", 32; Duff, *Punishment, Communication, and Community*, 139-141.

[332] Von Hirsch, *Censure and Sanctions*, 15, 25.

To illustrate my point, I wish to consider two cases. First, suppose we implement a policy giving welfare aid to disabled persons. The real basis of this policy may be that it is most politically feasible to help the disabled so we do it first. However, a lot of people may believe that the state is sending out a message that the abled body poor persons should be held responsible for their own poverty, which is not true. Second, suppose we used to hold lavish festivals for veterans (or other people performing heroic acts that should be encouraged) in recognition of their service. Now we decide that we should cut the festival budget and use the money for social welfare. A lot of people may believe that the state is sending out a message that the veterans are no longer respected, which need not be true. In both cases, our policy, which improves the well-being of the worst off members of society, may express wrong messages. But while that provides *a pro tanto reason* against those policies, it is plausible to believe that those policies *need not* be wrong, all things considered. Similarly, the mere fact that we may send out a wrong message by mitigating an offender's punishment in order to improve the well-being of the worst off does not show that generally we should not do so, all things considered. If anything, it is much easier to explain why the state, a coercive institution, should care about general distributive justice than to explain why it should care about the expression of true ideas.[333]

[333] For a defence of the idea that expression is a proper function of a liberal state, see Metz, "How to Reconcile Liberal Politics with Retributive Punishment".

I wish to consider two replies from my opponent. First, he may argue that actually no wrong message is *expressed* (as opposed to *perceived*) in the two cases above (the disabled welfare case and the veteran festival case) so those cases are not counterexamples to the claim that the importance of expressing correct message generally trumps the importance of improving the worst off. He may argue that one expresses a wrong message *only if he intends or aims to express that message.* In the two cases above, the wrong messages were no*t intended* but rather were merely *unfortunate side effects* of the policies.

This reply cannot assist my opponent. *If* one expresses a wrong message only if he aims to express that message, then it is not true that we would express a wrong message about the offender's blameworthiness for his crime by reducing his punishment in order to improve the well-being of the worst off. After all, we reduce the punishment *not because* we think the crime is less serious *but only because* we want to improve the worst off – ideally we would wish the public to understand the real basis of mitigation of sentence and we may even want to clarify this point as we hand out sentences. *If* there is *any* concern about expressing the wrong message about the offender's blameworthiness for his crime in mitigating his punishment for distributive justice, it must lie in the *effect* of doing so rather than the *aim* of the punisher.

The second line of reply to my point argues that while expressive concerns, in general, do not enjoy a priority over the importance of improving the worst off, the *particular expressive concern in punishment* is much more important. But what makes punishment special? There are three possible suggestions. First, the message being sent out in punishment is not any other ordinary message – it is one about the offender's blameworthiness for his crime. Second, the message being sent out in punishment implies something about the status of the victim. Third, punishment is explicitly expressive.

In reply, the first suggestion seems entirely *ad hoc*. There seems to be no reason why we should care more about the state's view on an offender's blameworthiness for his crime than the state's view on what kind of heroic conducts are encouraged or what the requirements of general distributive justice are.[334] As such, the mere fact that the message sent out in punishment is one about an offender's blameworthiness for his crime, in itself, does not explain why it is particularly important to express a correct

[334] Dolinko, "Some Thoughts about Retributivism", 551.

message about it, and accordingly, does not explain why expressing truths in punishment is more important than expressing true ideas in other contexts.

According to the second suggestion, if we mitigate the punishment of the offender, then we send out the message that the victim's interests count less.[335] It might be thought that avoiding wrong messages about the importance of a person's interests is more important than avoiding other kinds of wrong message, and accordingly, there is a special reason to express the correct message in punishment. The two cases above (the disabled welfare case and the veteran festival case) can then be distinguished on the ground that while false messages are sent out in those cases, the false messages in those cases do not involve the idea that someone's interests count less.

In reply, even if we accept that there is a special reason to avoid expressing a wrong message that someone's interests count less, I do not think that reason is sufficiently weighty to generally trump the importance of improving the worst off. Imagine that the blacks and the Latinos are the disadvantaged groups in society. Suppose we make a policy that helps the blacks but we do not make similar policies that help the Latinos because a policy that helps the Latinos would be politically infeasible or otherwise too

[335] For a similar point, see Lippke, "Social Deprivation as a Tempting Fate", 290.

costly. (Suppose further that our policy did not make the Latinos worse off.) Now even if our act has a legitimate moral basis, it is entirely possible that our act will be perceived as expressing the idea that the interests of the blacks count more than the interests of the Latinos. But that does not seem to be a decisive objection against the policy. Similarly, even if mitigating punishment for improving the well-being of the worst off has the unfortunate effect of misleading people into believing that the state regards the interests of some victims as counting less, it does not seem to me to be a decisive objection against it.

The third suggestion focuses, *not* on the special *content* of the message being sent out in punishment, but on the fact that punishment is an *explicitly expressive* act, i.e. an act that is explicitly stated to express a message. The idea is that if someone engages in an explicitly expressive act, then there is a special duty to send out the correct message. Since punishment is explicitly expressive, there is a special duty to send out the correct message in punishment.

I doubt, for two reasons, whether this suggestion can show that expressing the correct message in punishment has a general priority over improving the well-being of the worst off. First, I doubt if it is much more wrongful to express a wrong message in an explicitly expressive act than expressing a wrong message otherwise. Consider two cases. In the first case, the state explicitly tells a citizen that he is worthless. In the second case, the state does not explicitly tell him that. But the state's acts towards him clearly express the message that the state regards him to be worthless. Is the state's act much less wrongful in the second case? I think not. Perhaps there is a difference in wrongfulness but I do not think there is a big difference. Second, it seems that holding festivals for veterans is an explicitly expressive practice. So if sending out wrong messages for improving the well-being of the worst off is acceptable in such a case, then there seems to be no general bar against sacrificing expressive concerns for improving the well-being of the worst off even in cases involving explicitly expressive practices.

The second attempt to justify proportionality comes from the unfair advantage theory of punishment, which we have discussed in section 3 of this chapter above. According to this theory, if an offender is more blameworthy for his crime then he gains more unfair advantage.[336] As a result, a heavier punishment would be needed to remove his extra advantage. Three points can be made here. First, it is unclear how, even in a society that is distributively just, an offender must gain benefit by offending (see the

[336] Sher, *Desert,* 81.

unclear benefit objection above), and how the amount of benefit gained is proportional to the offender's blameworthiness for his crime. A murderer is clearly more blameworthy for his crime than an offender who committed tax evasion, but is it plausible to believe that the murderer must have gained more advantage?[337] Second, even if the unfair advantage theory of punishment can justify proportionate punishment in a society that is distributively just, it is, as Murphy has shown, hard to see how the idea of unfair advantage can justify proportionate punishment in the real world, which is distributively unjust.[338] Third, even if we set both points aside and assume that the unfair advantage theory of punishment can justify proportionate punishment in a distributively unjust world (perhaps by postulating *a kind of good* the gain of which is proportional to the offender's blameworthiness for his crime and independent of other goods like freedom or property such that even a socially deprived offender can have a net benefit of that kind of good after offending), we still lack an explanation as to why the removal of *this kind* of unfair advantage should generally take priority over improving the well-being of the worst off. The importance of removal of this kind of unfair advantage seems to lie in the importance of fairness. But the reason why we want to improve the well-being of the worst off members of society lies in fairness as well. Why is a fair distribution of *the kind of good that*

[337] Dolinko, "Some Thoughts about Retributivism", 545.
[338] Murphy, "Marxism and Retribution".

concerns the unfair advantage theory of punishment, whatever it is, more important than a fair distribution of *well-being*?

The third attempt to justify proportionality does not appeal to abstract principles like expressive concerns or unfair advantage. Rather, it simply claims that proportionality is one of our firmest considered judgments.[339] I have some sympathy with this approach to justify proportionality but if proportionality is grounded on this "non-foundationalist"[340] intuitive approach then I doubt, for two reasons, if we should believe proportionality to have a general priority over improving the well-being of the worst off. First, while many people have the intuition that there is a reason to make punishment proportionate, it is unclear to me that most people have the intuition that making punishment proportionate is the first virtue in punishment. Second, I suspect most people have a strong intuition that improving the worst off is an important moral concern as well.

Here, my opponent may raise the following thought experiment. Suppose that by imposing the death penalty on a petty thief, we can improve the well-being of the worst off members of society greatly. Most people believe that we should not do so all

[339] Moore, "The Moral Worth of Retribution".
[340] Ryberg, *The Ethics of Proportionate Punishment,* 44.

things considered. Doesn't it show that, generally, we should not compromise proportionality for improving the well-being of the worst off?

In reply, I do not think that case shows that proportionality generally trumps the value of improving the well-being of the worst off because while proportionality sets both upper and lower limits to punishment, as I mentioned in chapter 1, I believe that there is an asymmetry between the strength of them and going above the upper limit is much harder to justify than going below the lower limit. It is plausible to believe that, given the importance of treating offenders as ends, there are strong *upper* limits to punishment set by offenders' blameworthiness for their crimes. That explains why imposing the death penalty on a petty thief is generally unacceptable, all things considered, no matter the consequences. However, the importance of treating an offender as an ends does not provide any reason against *downward* deviation from proportionality, and hence, does not provide any reason for believing that generally speaking, we should not, all things considered, *mitigate* an offender's punishment to improve the well-being of the worst off members of society.[341]

When we reflect on some commonly accepted practices in our sentencing system, we

[341] Here I am agreeing with Ryberg, *The Ethics of Proportionate Punishment*, 48.

should be even more sceptical towards the claim that proportionality is a lexically trumping value in punishment. For example, we should remind ourselves that the most important mitigating plea in our criminal justice system is one that violates proportionality: the mitigation given to offenders who plead guilty.[342] The guilty plea conflicts head on with proportionality as in essence this plea entails that two equally blameworthy offenders can attract unequal punishment depending on whether the offender pleads guilty.[343] Unless we want to say that we should abolish the guilty plea, we implicitly accept that punishment is a practice in which a plurality of considerations comes into play, of which proportionality may be just one major player amongst many others.[344]

The third argument for claiming that generally we should not mitigate an offender's punishment to improve the well-being of the worst off members of society, all things considered, appeals to the importance of *crime-reduction.* According to this argument, if we mitigate an offender's punishment to improve the well-being of the worst off members of society then more crimes will be committed, and generally speaking

[342] The guilty plea is recognized all around common law jurisdictions. For England, see *Criminal Justice Act 2003* s114.

[343] To be fair, some prominent writers, like Ashworth, seem to have doubts about whether we should recognize this plea. See Ashworth, *Sentencing and Criminal Justice: fifth edition*, 178-180. But his reasons for rejecting the guilty plea have nothing to do with proportionality and thus are not directly relevant here. Many writers agree with me on the point that no single value dominates punishment. See, e.g., Cahill, "Punishment Pluralism"; Berman, "Two Kinds of Retributivism".

[344] See also Brown, "Third Party Interests in Criminal Law".

crime reduction is more important than improving the well-being of the worst off members of society.

There are two main problems with this argument. First, in cases where crime reduction conflicts with the improvement of the well-being of the worst off, it is unclear why we should generally favour the former. The central reason, if not the sole reason, why crimes ought to be prevented is because they harm people's well-being. As such, the main reason why we want to minimize crimes is that we want to maximize well-being. As I mentioned above, there is no reason why maximization of well-being is generally more important than improving the well-being of the worst off.[345]

The second, and more important, problem of this argument is that there is no reason to believe that in a reasonably just sentencing system (i.e. a sentencing system that is constituted by sentences that are appropriate *before* we consider mitigation based on improving the well-being of the worst off), a reduction of an offender's punishment *must always, or nearly always*, result in a substantial increase in crimes. Hence, there is no reason to believe that the concern for crime reduction, even if paramount,

[345] Rawls, *A Theory of Justice: revised edition.*

supports the claim that generally we should not, all things considered, mitigate an offender's punishment in order to improve the well-being of the worst off. Let me be very clear about what I am arguing for here. I am *not* arguing that mitigation of sentence (in a reasonably just sentencing system) will *never* increase crimes substantially. I am only arguing that *in many cases* it will not. That is sufficient for my purpose, since I only aim to argue that in many cases, all things considered, we should reduce an offender's punishment if that will improve the well-being of the worst off; I *never* intend to argue that we should *always* do so, all things considered.

If the gravity of sentences in a reasonably just sentencing system is solely set by crime reduction concerns (i.e. for each type of sentence, we decrease it until crime rate significantly increases) then there may be reasons to believe that any mitigation of sentence will always increase crimes substantially. However, as I mentioned in chapter 1, in a reasonably just sentencing system the sentences will not be set solely by crime reduction concerns but will be set partly by proportionality. It is hard to believe that reducing a proportionate sentence *must always or nearly always* increase crimes.

I believe my position is supported, both by theoretical concerns and also indirectly by the existing empirical researches on the relationship between crime rate and the severity of sentences. Let me start with the theoretical concerns. The main pathways through which crimes may be reduced by an increase in the severity of sentences are, according to most writers, deterrence and incapacitation. It is implausible to believe that reducing a proportionate sentence will always or nearly always substantially compromise deterrence or incapacitation.

A decrease in sentence severity will only substantially reduce the deterrent effect if the difference in penalty is significant for a lot of potential offenders and if a lot of potential offenders think the possibility of apprehension is substantial. For some types of crime, the perceived possibility of apprehension is low so reducing sentences need not substantially reduce the deterrent effect.[346] For some crimes, a lot of times they are committed by people who act under impulse.[347] Moreover, the significance of mitigating a certain piece of sentence, in the mind of potential offenders, may not be great if the post-mitigated sentence is still heavy. To illustrate, the difference between absolute discharge and a two-year sentence may be an important consideration for many potential offenders; but the difference between a five-year sentence and a seven-year sentence (or a nine-year sentence for this matter) may not be a significant

[346] Ashworth, *Sentencing and Criminal Justice: fifth edition*, 80.
[347] UK Home Office, *Crime, Justice, and Protecting the Public*.

consideration for many offenders. This is so for two reasons. First, for many offenders, they will not think in very precise terms: they will just think about whether the punishment is "heavy" or not. If the post-mitigated sentence is still heavy, a change from a very heavy sentence to a less heavy but still very heavy sentence will not affect their reasoning much. Second, many people will have difficulty imagining what would happen several years down the road, and as a psychological matter people will discount the future if they do not imagine it vividly.[348] Even if sentences in a reasonably just sentencing system will not be as heavy as the ones that currently exist in the US, it does seem plausible to believe that it will contain a number of long sentences for serious crimes like murder and rape. It follows that for many sentences in a reasonably just sentencing system, reducing them is unlikely to substantially compromise the deterrent effect, even if the possibility of apprehension for those crimes is not small.

Incapacitation can only substantially reduce crimes if the replacement effect (i.e. the crime that would be committed by the prisoner had he been free is committed by someone else) is small and if the prisoner would commit a lot of crimes if let free. Given that some crimes have great replacement effect (e.g. it is well documented that

[348] Bauer, "How to Avoid the Temptations of Immediate Gratification".

if we incapacitate drug traffickers, it is unlikely to result in much reduction in drug trafficking – since the big boss would just hire others to do the job) and we know that most criminals reduce or stop their criminal behaviour as they grow old,[349] a longer sentence can only substantially reduce crimes through incapacitation for certain types of crime and for certain types of offender and accordingly there is no reason to believe that a reduction of sentence will generally diminish incapacitation effect substantially.[350]

Empirical researches on the relationship between crime rate and severity of sentences also provide indirect support for my position. Before we look at them, let me stress at the outset that I believe the support they offer for my position is only indirect. It is because the question I am asking is whether a reduction of sentence, *in a reasonably just sentencing system*, must always or nearly always increase crimes. However, the empirical studies on the relationship between crime rate and severity of sentences are done on *existing sentencing systems*, and the sentences in our existing systems need not be reasonably just. Accordingly, any support such researches can provide for my position can only be indirect. That said, I believe they can offer *some* support for my position – I will elaborate on this point after I go through the empirical researches.

---

349 Piquero and Blumstein, "Does Incapacitation Reduce Crime".
350 Piquero and Blumstein, "Does Incapacitation Reduce Crime"; Tonry, "Sentencing in America: 1975-2025", 181-182.

In the literature, there are optimists and pessimists about how much extra crime reduction effect we can get by increasing our existing sentences (or how much crime reduction effect we will lose by reducing existing sentences) within a reasonable level (i.e. we are not talking about changes like adding twenty years to petty theft or abolishing the existing criminal justice system altogether).[351] Some scholars, like Doob and Webster, are more pessimistic. In 2003, after reviewing a large number of previous empirical studies, they concluded that an increase in sentence severity in present America, within certain limits, generally would not reduce crimes.[352] A similar conclusion was drawn by Tonry in a 2013 paper after examining previous government reviews for policy making, scholarly surveys, and evaluations of mandatory minimum and three-strikes laws.[353] Many studies done outside the US reported similar results.[354]

Other scholars are more optimistic about the crime reduction effect of increasing the severity of sentences. Abrams, in a 2011 paper, argues that add-on gun laws (i.e. laws that "stipulate sentence enhancements for defendants convicted of possessing of a

---

351 Doob and Webster, "Sentence Severity and Crime: Accepting the Null Hypothesis", 146 n1.
352 Doob and Webster, "Sentence Severity and Crime: Accepting the Null Hypothesis".
353 Tonry, "Sentencing in America 1975-2025".
354 See the ones discussed in Tonry, "Sentencing in America 1975-2025", 177-178.

firearm while committing a crime") provided a significant deterrent effect on gun robberies (interestingly though in his study he fails to find a significant deterrent effect of such laws on gun assaults).[355] Similarly, Bell et. al., in their recent paper, argue that an increase in the severity of sentences sometimes substantially reduce crimes by comparing crime rates before and after the harsh sentences given out during the London Riots.[356] Unfortunately, the studies by Abrams and Bell et. al. were not discussed by Tonry.

What conclusion can we draw from the empirical researches? I think the dispute between the optimists and the pessimists is one about whether it is true that crime rate is *generally insensitive* to changes in the severity of sentences (within certain limits) in societies similar to ours: the pessimists think the answer is yes while the optimists think the answer is no. I do not want to take a side in that dispute. Instead, I want to draw a weaker conclusion, which is sufficient for my purpose and which I think will be accepted by both sides. I believe that even the optimists would not necessarily disagree with the claim that while an increase in the severity of sentences *can* substantially reduce crimes *in some cases*, it will only do so in certain favourable circumstances and *will not do so in many other cases*. In fact, even Abrams, an

355 Abrams, "Estimating the Deterrent Effect of Incarceration using Sentencing Enhancements".
356 Bell et. al., "Crime Deterrence: Evidence From the London 2011 Riots".

optimist, only found a correlation between crime rate and severity of sentence for some types of crime (like gun robberies) and *failed to find that for other types of crime* (like gun assaults).

Of course, the claim that I am arguing for is "a reduction of punishment of some offenders will not (nearly) always increase crimes in *a reasonably just sentencing system*" rather than the claim that "a reduction of punishment of some offenders will not (nearly) always increase crimes in *our existing sentencing systems*", so the empirical researches I just discussed do not provide direct support for my claim. That said, there is an important similarity between existing sentencing systems and a sentencing system that is reasonably just: they are not entirely structured by crime reduction concerns and are partly structured by proportionality. So even if the existing empirical researches do not conclusively prove that a reduction of sentence will, in many cases, not increase crimes in a reasonably just sentencing system, I believe the burden of proof rests on my opponent should he wish to claim otherwise.

The fourth argument for claiming that generally we should not mitigate an offender's punishment in order to improve the well-being of the worst off members of society, all things considered, appeals to the importance of rehabilitation. This argument claims that if we mitigate an offender's punishment to improve the well-being of the worst off members of society, then we reduce the chance of securing the offender's rehabilitation, a loss that morally outweighs the value of improving the well-being of the worst off.[357]

There are two main problems with this argument. The first problem is that it is unclear why securing the reform of the offender is generally more important than improving the well-being of the worst off. In fact, there is an indirect argument for claiming otherwise. Many people believe that the importance of securing the reform of the offender does not trump proportionality. Suppose by giving a brutal rapist an unusually short sentence (say one day of jail time) we are more likely to secure his moral reform than a normal sentence. Many people are against such reduction of sentence on the ground that it violates proportionality. But if the importance of securing the offender's moral reform does not trump proportionality, and as I have defended above that proportionality does not trump the value of improving the well-being of the worst off, then there is no reason to believe that the importance of securing the reform of the offender should generally trump the value of improving the well-being of the worst off.

[357] While this argument appeals to Duff's idea of secular penance, it should be noted that Duff himself justifies proportionality on communicative ground instead. See Duff, *Punishment, Communication, and Community*, 132.

The second main problem of this argument is that it is unclear that a decrease in the length of sentence in the real world, within certain limits, will generally have much negative effect on the rehabilitation of the offender.[358]

The fifth argument for why, all things considered, we should not use sentencing to improve the well-being of the worst off members of society rests on the idea of judicial competence. In order to know whether mitigating the punishment of an offender will have that effect, we need to know the broader social consequences of doing so (e.g. we need to know whether punishing the offender less will increase crime rate). But how is it possible for judges to do that kind of consequential calculus in sentencing? “Judges are not trained economists.”[359]

In reply, my opponent’s argument assumes that since sentences are given out by the judges, the consequential calculus must be done by the judges as well. But that is not the case. The legislature can find out for which categories of offenders reducing their

---

[358] If anything, most studies on recidivism show that an increase in the severity of sentence has no effect on recidivism or increases recidivism for many offenders. For some surveys, see, e.g., Song and Lieb, *Recidivism: The Effect of Incarceration and Length of Time Served*; Lipsey and Cullen, “The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews”; Nagin, Cullin, and Jonson, “Imprisonment and Reoffending”.

[359] Burke, “Reinventing Contracts”. For doubts as to judicial competence in taking into account consequentialist considerations in decisions, see, e.g., Sunstein, “If People Would Be Outraged by Their Rulings, Should Judges Care?”; Brodeur and Roberts, “Taking Justice Seriously”; Vincent, “Book Review of *Torts, Egalitarianism and Distributive Justice*”; von Hirsch, *Doing Justice,* 146.

punishment will improve the well-being of the worst off members of society and issue guidelines to judges. So what the judges need to do in sentencing an offender is merely to verify whether the offender falls into those categories. That task would be well within the judges' competence.[360]

The sixth argument for why, all things considered, we should not use mitigation of punishment to improve the well-being of the worst off members of society claims that even if the task of getting information about the broader social consequences of mitigation of punishment is assigned to the legislature rather than the judges, such a task is too difficult and costly.

In reply, to acquire infallible or perfect information about the consequences of mitigation of sentence would of course be very expensive. But perfect information, while desirable, is not necessary. Lack of perfect information would prevent us from doing perfect justice; but that does not seem to be a strong argument for doing nothing. In most cases where we try to achieve some good with a social policy, say an education policy, we act under a certain degree of ignorance and fallibility.[361] And it

[360] For a suggestion of a reply alone this line, see Vincent, "Book Review of T*orts, Egalitarianism and Distributive Justice*".

[361] I may have read this reply somewhere but I am afraid I cannot trace the exact source at this moment.

is unduly pessimistic to think that we cannot acquire reasonably reliable, even if imperfect, information about the broader social consequences of sentencing policies in the foreseeable future. For example, recently there have been a number of empirical studies conducted on the broader social consequences of mass incarceration in America and I see no reason why it is unrealistic to conduct some on the broader social consequences of other sentencing policies in the foreseeable future.[362]

5. The empirical claim of the improvement of the worst off argument

Let me now move on to the empirical claim of the improvement of the worst off argument, the claim that for a greater proportion of socially deprived offenders than non-socially deprived offenders, reducing their punishment will have the consequence of improving the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices).

Before examining the empirical claim, I wish to remind the readers that the empirical claim is *not* the claim that *all or nearly all* socially deprived offenders satisfy the following description: "reducing this offender's punishment will improve the

[362] See, e.g. Western and Pettit, *Collateral Costs: Incarceration's Effect on Economic Mobility*.

well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices)". As explained in chapter 1 of this thesis, that will set too high a bar for a successful argument. The empirical claim is true as long as the proportion of socially deprived offenders who satisfy that description is *higher* than that for those who are non-socially deprived offenders. For example, if 30% of socially deprived offenders satisfy that description while only 5% of non-socially deprived offenders satisfy it, then the empirical claim is true.

Despite that reminder, I believe that without further empirical evidence, we are *not entitled* to accept the empirical claim of the improvement of worst off argument. This may sound surprising to some because, as I mentioned in section 1 of this chapter, there seems to be a simple intuitive line of reasoning for the claim. First, it seems plausible to believe that the disadvantage of most worst off members of society are not justified by their responsible choices (as mentioned in section 1 of this chapter, I will assume this). Second, socially deprived offenders are the worst off members of society, while that is not the case for non-socially deprived offenders. Third, reducing the punishment of a socially deprived offender improves the well-being of him, a worst off member of society; accordingly, mitigation of punishment of a socially deprived offender will improve the overall well-being of the worst off members of society. Fourth, reducing the punishment of a non-socially deprived offender generally will only benefit him, who is not a worst off member of society; accordingly mitigation of punishment of a non-socially deprived offender generally will not improve the well-being of any worst off members of society.

However, I believe that the third step and the fourth step of the above chain of reasoning are not solid once we take into account the *indirect* effects of mitigation of punishment. I want to use an analogy to illustrate my point. The belief that increasing welfare payment for the poor will benefit the poor is intuitive. Such belief is usually based on the idea that the direct effect of welfare payment is beneficial to the poor. But that would be oversimplifying the matter as we need to take the indirect effects of welfare payment, which may be counterproductive, into account. Increasing welfare payment may harm capitalistic activities, which may harm the poor in the long run. Whether increasing welfare payment will ultimately improve the position of poor people is a complicated empirical question. Analogously, I believe there are three considerations looking at the *indirect* effect of mitigation of punishment that, *adding together,* cast doubt on the empirical claim of the improvement of the worst off argument.

First, while I have argued in section 4 of this chapter above that, *for many* socially deprived offenders, reducing their punishment will improve their own well-being, it should also be observed that *for some other* socially deprived offenders, reducing their punishment will not have that effect. While the direct effects of mitigation of punishment, i.e. increase in freedom or money, generally promote well-being, the indirect effects of mitigation of punishment *need not* be. There are some empirical studies showing that, for many offenders, an increase in the length of sentence will increase the income and employability of them at least in the short term.[363] While noting that such an effect does not obtain for many other offenders and income does not equate with well-being such that a person who spent more years in jail may have less well-being in total even if his future income is slightly increased by the increase in sentence, it seems safe to conclude that for *some* socially deprived offenders, reducing their punishment will not improve their own well-being.

Second, for some socially deprived offenders, reducing their punishment may harm *other* worst off members of society; accordingly, even if mitigating their punishment will improve their own well-being such a measure will not improve the overall

[363] Jung, "Increase in the Length of Incarceration and the Subsequent Labor Market Outcomes: Evidence from Men Released from Illinois State Prisons"; Landerso, *Does Incarceration Length Affect Labor Market Outcomes for Violent Offenders?*; Kling, "Incarceration Length, Employment and Earnings".

well-being of the worst off members of society. For some offenders, their families will be better off if they stay in jail.[364] More importantly, while mitigation of punishment does not, as I have argued in section 4 of this chapter above, always increase crimes, it certainly does *sometimes.* Since most of the victims of crimes are also worst off members of society,[365] if by mitigating the punishment of a socially deprived offender we increase crimes then we may not improve the overall well-being of the worst off members of society even if we benefit the offender.

Let me be clear about what these two considerations are supposed to show. These two considerations do *not* deny that for *many, or a significant proportion of*, socially deprived offenders, reducing their punishment will improve their own well-being without harming other worst off members of society, and accordingly, have the effect of improving the overall well-being of the worst off members of society. In fact, given my discussion in section 4, I believe that is true. These two considerations *only* show that it is not true that *for all or nearly all* socially deprived offenders, reducing their punishment will have that effect. As such, these two considerations, *per se*, would *not* pose a threat to the empirical claim of the improvement of the worst off argument *if* generally speaking reducing the punishment of a non-socially deprived offender will

[364] See a tentative study, see Wakefield and Wilderman, *Children of the Prison Boom*, chapter 4.
[365] Karman, "Poverty, Crime, and Criminal Justice", 31-34.

not improve the well-being of any worst off members of society. However, the third consideration precisely provides reasons to doubt that.

The third consideration is that reducing the punishment of a non-socially deprived offender may, in many cases, have the indirect effect of improving the overall well-being of the worst off members of society. This may sound surprising: how can lessening a burden on the rich help the poor? We can, of course, imagine *exceptional* cases where mitigating the punishment of a non-socially deprived offender will benefit the worst off members of society. For example, we can imagine a rich factory owner, who hires a lot of poor people, will be forced to close his factory if a lengthy prison term is imposed on him. But why should we believe that such a case is anything other than exceptional; why should we believe that mitigating the punishment of non-socially deprived offenders may have that effect in many cases?

The answer lies in that, in many cases, we can use reduction of sentences on some non-socially deprived offenders *to create incentives to encourage behaviours that will improve the well-being of the worst off*. To illustrate, think about the following policy: we reduce the punishment of some offenders who donated a certain amount of money to charity (or offenders who are family members of people who donated money to charity) by a small amount.[366] This mitigating policy may improve the overall well-being of the worst off members of society (as it will encourage people to donate) if such a policy does not increase crimes or otherwise harm distributive justice, and may cover a lot of non-socially deprived offenders. I do not think we have data as to what proportion of non-socially deprived offenders would have a mitigating plea under such policies; there seems to be no basis to assume that the proportion must be insignificant.

I do *not* take these three considerations to have conclusively established the falsity of the empirical claim of the improvement of the worst off argument. But they cast doubt on the claim: the first two considerations show that it is not true that for *all* socially deprived offenders, mitigating their punishment will improve the overall well-being of the worst off members of society as we will obtain this effect for only some of these offenders; the third consideration shows that for a portion of non-socially deprived offenders, mitigating their punishment may also have that effect. As such, I think we are not entitled to commit to the empirical claim of the argument without further empirical data.

[366] For a discussion of related policies, see Ashworth, *Sentencing and Criminal Justice: 5th ed*, 182.

I wish to conclude the discussion of the empirical claim by reemphasizing one point. I have only doubted whether the mitigating factor based on improving the well-being of the worst off members of society has a *greater* applicability to socially deprived offenders than to non-socially deprived offenders; I have *not* doubted that this mitigating factor is applicable to *many* socially deprived offenders (as I said, I actually believe that is the case).

6. Will socially deprived offenders fare worse, all things considered, under a sentencing system that takes into account the impact of sentence on the worst off members of society, as compared to a sentencing system that does not?

In the above I have argued that the idea of improving the worst off members of society should be taken into account in sentencing and this idea can justify mitigation of sentence for many socially deprived offenders (even if the idea may also justify mitigation of sentence for many non-socially deprived offenders). It may be objected that, nonetheless, taking into account the impact of sentence on the worst off members of society in sentencing will harm socially deprived offenders all things considered, as compared to a sentencing system that does not take it into account. This is because socially deprived offenders are often *more dangerous*, so aggravating their punishment can reduce crimes, and reducing crimes will benefit the worst off members of society; it is actually better for the socially deprived offenders if we simply exclude all consequentialist calculus from sentencing.[367]

I have two replies. First, to take into account the impact of a sentence on the worst off members of society through *mitigation* does not mean that we have to take it into account through *aggravation*. I have provided the reasons above in section 4 of this chapter and chapter 1 of this thesis: aggravation of sentences in a reasonably just sentencing system often involves punishing an offender more than what he deserves based on proportionality; and punishing an offender more than what he deserves based on proportionality involves treating him as a mere means, which is not involved in punishing him less than what he deserves based on proportionality. Accordingly, it is entirely consistent for me to claim, on the one hand, that we should, all things considered, mitigate the punishment of an offender for distributive justice in many cases, while on the other hand, generally we should not, all things considered, aggravate the punishment of an offender to promote distributive justice.

Second, as an empirical matter, crime rate is not always sensitive to the severity of

[367] Von Hirsch, *Censure and Sanctions*, 99; Hudson, "Beyond Proportionate Punishment", 67.

sentences so in many cases increasing the severity of sentences will not reduce crimes. I have discussed this point in section 4 of this chapter above.

7. Robustness in relation to theories of distributive justice

In section 2.2 of this chapter above, I have defended a particular conception of general distributive justice, namely, responsibility-catering well-being prioritarianism. My discussion on the improvement of the worst off argument proceeds on that conception of general distributive justice. Of course, I am aware that the question of what are the requirements of general distributive justice is one of the most controversial topics in political philosophy, and I do not expect every reader to accept my view on that topic. In this section, I will explore to what extent my discussion concerning the improvement of the worst off argument is robust in relation to different theories of distributive justice, that is, to what extent my main claims concerning the improvement of the worst off argument, and my defence of those claims, can remain intact (or only require minor amendments) if my conception of distributive justice is given up. Unsurprisingly, my discussion on the improvement of the worst off argument will survive less radical disagreements on distributive justice but will be inconsistent with views on distributive justice that are more radically different from mine.

Let me first recap the four main claims I made concerning the improvement of the worst off argument:

Claim (1): A measure that improves the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices) promotes distributive justice to a greater extent than a measure that does not.

Claim (2): If by reducing the punishment of an offender we will promote distributive justice to a greater extent than mitigating the punishment of another offender, then there is an extra reason to reduce the punishment of the former.

Claim (3): Not only do we have an extra reason to reduce the punishment of an offender if by doing so we will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices); in fact, in many cases we should do so, all things considered.

Claim (4): For many (but not all) socially deprived offenders, reducing their punishment will improve the well-being of the worst off members of society (whose disadvantage is not justified by their responsible choices); however, for many non-socially deprived offenders, reducing their punishment may also have that effect.

Let us start with people who disagree with me only as to the *currency* of general distributive justice. I believe that the proper currency of general distributive justice is well-being, but I do not think accepting something else, like primary goods, as the proper currency of general distributive justice would affect my discussion much. We only need to substitute the phrase “well-being” in the four original claims with “primary goods” to form four amended claims (I will label the amended claims as claims (1’), (2’), (3’), and (4’) respectively), which are similar to the original claims (the two sets of claim are similar as people who have less well-being are likely to be those who have fewer primary goods and vice versa); and I believe the defence I offered for claims (1)-(4) will largely apply to justify claims (1’) to (4’).

To illustrate, I defended claim (1) by appealing to the idea that unchosen disadvantage is morally arbitrary and hence unjust; it seems claim (1’) can be defended in a similar manner. Claim (2’) is identical to claim (2) so my defence of claim (2) applies automatically to claim (2’). I defended claim (3) by comparing, for example, the value of improving the position of the worst off members of society with the value of proportionality. My main arguments and examples there do not depend on a very particular conception of “worst off members of society”. For example, the veteran festival example seems to show that sometimes expressive concerns can be sacrificed for improving the position of the worst off members of society, no matter whether the “worst off” is ultimately defined in terms of well-being or primary goods. In relation to claim (4’), while conceptually speaking primary goods and well-being are different, practically speaking the group that consists of worst off members of terms of well-being is likely to be largely co-extensive with the group that consists of the worst off members of terms of primary goods; accordingly, my defence of claim (4) will largely apply to justify claim (4’).

In addition to being largely robust towards different views about the *currency* of general distributive justice, I believe my discussion concerning the improvement of the worst off argument is also robust towards some (but not all) views about the proper *principle* of general distributive justice that are different from mine.

I believe that the correct principle of general distributive justice is responsibility-catering, but giving up that belief will only necessitate a minor amendment to my main claims concerning the improvement of the worst off argument. All we need to do is to take out the phrase “(whose disadvantage is not justified by their responsible choices)” from claims (1) to (4) to form claims (1’’) to (4’’), which are similar. I believe the defence I offered for claims (1) to (4) largely applies to justify claims (I’’) to (4’’). For example, in relation to claim (I’’), *if* it is true that responsible choices do not matter to general distributive justice, then it follows that any disadvantage, chosen or not, is morally arbitrary and hence unjust and accordingly we should accept a version of prioritarianism that does not take into account responsible choices. Claim (2’’) is identical to claim (2). I believe my defence of claim (3) can largely apply to justify claim (3’’) as well: for example, the gist of the veteran festival example is to show that sometimes it is acceptable to use goods that can otherwise be used for expressive concerns to help the worst off members of society; the plausibility of the example does not seem to be closely tied to the issue whether the worst off members of society are responsible for their own disadvantage. Since the disadvantage of most worst off members of society cannot be justified by their responsible choices, my defence of claim (4) will largely apply to justify claim (4’’) as well.

I believe that the correct principle of general distributive justice is prioritarian, but I believe an egalitarian or a sufficientarian will have no problem accepting claims (1) to (4) and my defence of them. An egalitarian and a sufficientarian will disagree with me as to whether distributive justice will be improved by giving more goods to someone who unjustifiably has a lot of goods. But they will agree with me that the most urgent moral concern is to improve the position of the worst off members of society who are below the line of sufficiency, and accordingly, they can accept claim (1); claims (2) to (4) and my defence of them can clearly be accepted by an egalitarian or a sufficientarian.

In the above, I have not simply subscribed to prioritarianism as a principle of general distributive justice; I have also subscribed to the idea that the priority is *absolute,* insofar as general distributive justice is concerned (i.e. I have subscribed to a *maximin* version of prioritarianism as a theory of general distributive justice). How will my discussion of the improvement of the worst off argument be affected if we accept a non-maximin version of prioritarianism, i.e. a theory of general distributive justice that assigns some, but not absolute, weight to the well-being of worse off?

It depends on how much weight this prioritarian theory gives to the worse off. Let us form four amended claims, claims (1’’’) to (4’’’) corresponding to this theory. (For example, claim (1’’’) reads “A measure that increases weighted utility to a greater extent promotes distributive justice to a greater extent”.) If the weight is a strong one, then a measure that improves the well-being of the worst off is likely to be a measure that increases weighted utility to a greater extent, and accordingly, claims (1) to (4) will be largely similar to claims (1’’’) to (4’’’) and my defence of the former set of claims will largely apply to the latter set. However, if the weight is a small one, then claims (1’’’) to (4’’’) will be very different from claims (1) to (4); accordingly my discussion above is not robust in relation to a prioritarian theory that assigns only a small weight to the well-being of the worse off.

Claims (1) to (4) and my defence of them would not be acceptable to someone who disagrees with me on the requirements of distributive justice in a more fundamental manner. A utilitarian will reject claim (1) and any similar claims since he would not see any special urgency to improve the position of someone who has the smallest amount of general goods. Libertarianism, Anderson’s democratic equality theory and the purely localist conception of distributive justice will also reject claim (1) and any similar claims since they do not accept that justice requires a particular distribution of general goods. I hope that, in section 2.2 of this chapter above, I have convinced the readers that those views on general distributive justice are mistaken.

8. Consistency of my discussion with retributivism and crime reduction theories of punishment

Before I conclude this chapter, I wish to briefly observe to what extent my discussion above is consistent with retributivism and crime reduction theories of punishment, since it is a question I sometimes receive in relation to this chapter.

As to the question of whether my discussion is consistent with retributivism, it depends on what one means by "retributivism". Let us distinguish between two conceptions of retributivism.[368] (Both of them, let me stress, are versions of positive retributivism and are not versions of negative retributivism/side-constrained consequentialism.[369]) According to the first conception, which can be called *pro tanto retributivism*,[370] there is a *pro tanto* reason to inflict punishment on an offender for his crime that is proportionate to his blameworthiness for that crime. According to the second and much stronger conception, which can be called *all things considered retributivism*, not only do we have a reason to punish proportionately; in addition, in general, proportionality is the only consideration we should take into account in

[368] For similar ideas, see Husak, "Retributivism in Extremis", 12-13.
[369] For the idea of negative retributivism, see Duff, *Punishment, Communication, and Community*, 11.
[370] For the term "pro tanto retributivism", see White, "Pro Tanto Retributivism".

deciding the gravity of punishment, all things considered.

My discussion is consistent with *pro tanto* retributivism. In fact, I have stated in section 4 of this chapter above that a reasonably just sentencing system is partly structured by proportionality, and in stating that I have committed myself to the truth of *pro tanto* retributivism. I have explained in chapter 1 of my thesis that I will assume the truth of *pro tanto* retributivism in my thesis.

However, I have claimed that sometimes we should reduce a proportionate sentence for distributive justice, all things considered, and this claim is inconsistent with all things considered retributivism. I have provided arguments for rejecting all things considered retributivism in section 4 of this chapter above and I hope the readers found them convincing.

Similarly, as to the question of whether my discussion in this chapter is consistent with crime reduction theories of punishment, it depends on what we mean by those theories. We can distinguish between *pro tanto* crime reduction theory and all things considered crime reduction theory. My discussion is entirely consistent with a *pro tanto* crime reduction theory. However, since I have argued that crime reduction does not generally trump general distributive justice (in case the two ideals conflict) and general distributive justice, a consideration distinct from crime reduction, should, in many cases, be taken into account in deciding the proper level of punishment all things considered, I reject the all things considered crime reduction theory. The arguments for rejecting the all things considered crime reduction theory are again provided in section 4 of this chapter above.

9. Conclusion

In this chapter, I have explored the improvement of the worst off argument.

I have argued for the moral claim of the improvement of the worst off argument, the claim that if the well-being of the worst off members of society will be improved by reducing the punishment of an offender then there is a reason to reduce his punishment that does not apply to offenders the reduction of punishment of whom will not lead to that consequence, by appealing to the idea of general distributive justice. I have also argued that in many cases, all things considered, we should reduce an offender's punishment to improve the well-being of the worst off members of society.

I have, however, expressed reservations about the empirical claim of the improvement of the worst off argument, the claim that for a greater proportion of socially deprived offenders than non-socially deprived offenders, reducing their punishment will produce the consequence of improving the well-being of the worst off members of society. This is because reducing the punishment of some socially deprived offenders will not lead to that consequence, and reducing the punishment of many non-socially deprived offenders may lead to that consequence. Without further empirical data, therefore, I believe we are not entitled to accept the empirical claim of the argument.

While, as just mentioned, I think that the improvement of the worst off argument is not conclusively established, I believe my discussion of the argument leads to two important points. First, many writers hold that sentencing and distributive justice belong to distinct spheres, so even if we can promote general distributive justice by mitigating the punishment of an offender, there is no special reason to mitigate his punishment, or that generally we should not do so all things considered. In none of the major Western jurisdictions do judges mitigate the punishment of offenders on the ground that doing so will improve the well-being of the worst off members of society. I have challenged that view in defending the moral claim of the argument.

Second, although I have doubted the claim that the proportion of socially deprived offenders who can rely on the consideration of improving the well-being of the worst off members of society as a ground for mitigation of punishment is *greater* than that for those who are non-socially deprived offenders, nothing I said cast doubt on the claim that for many, or a significant proportion of, socially deprived offenders, reducing their punishment will improve their own well-being without harming other worst off members of society (in fact, I have put forward some reasons to accept this claim in section 4 of this chapter above). Accordingly, nothing I said cast doubt on the claim that many socially deprived offenders can rely on that consideration, a consideration that is not currently taken into account in sentencing practice, as a ground to mitigate their punishment. Therefore, even if the improvement of the worst off argument cannot establish its empirical claim, what it has established may still have profound influence on sentencing of socially deprived offenders in the present distributively unjust society.

Chapter 9: Moral education and burdens

1. Introduction

Some people believe that a *history* of bad moral education or upbringing is *intrinsically* significant to criminal punishment. In other words, they believe that there is a reason to punish an offender with a bad moral education less for his crime than an otherwise similar offender who had a good moral education, *even if they were identical at the time of offence*. This concern based on moral education is, let me stress, *distinct from the concern based on suffering or general distributive justice*. To illustrate, some people think that there is a reason to punish the son of a rich Mafia boss more leniently for his assault because of his bad moral education *even if* his total amount of general goods (welfare, primary goods, or resources) is not less than any other offender.[371]

There are two conceptually possible ways in which a factor, like a bad moral education, can justify punishing an offender less for his crime. First, the factor may be

[371] For the distinction between arguments based on suffering and arguments based on bad moral education, see Vuoso, "Background, Responsibility, and Excuse", 1665 n9.

relevant because it exculpates the offender, i.e. it reduced how much *blame*, i.e. negative reactive attitudes, the offender *deserves* for his crime. Second, the factor may be relevant in a non-exculpatory way, i.e. it provides a reason to reduce punishment *even if it does not affect the offender's blameworthiness for his crime.*

I have argued, in chapter 4 above, that an offender who had a bad moral education but was identical at the time of offence is *not less directly blameworthy* for his crime. (I have granted in chapter 3 that he is likely to be less indirectly blameworthy for his crime, but I have also argued, in that chapter, that a person's indirect blameworthiness for his crime should not be taken into account in deciding the punishment *for that crime*.)

But I have *not yet* examined whether moral education can be relevant to how much punishment we should impose on an offender for his crime *without affecting his blameworthiness for the crime*. It is conceptually possible that bad moral education can justify mitigation of punishment without diminishing blame; in fact, Paul Litton precisely argues so.[372] Rejecting the idea that bad moral education has such non-exculpatory significance to punishment is the main task of this chapter. Let me

[372] Litton, "The 'Abuse Excuse' in Capital Sentencing Trials", 1060-1066. See also Nussbaum, "Equity and Mercy".

lay out the two main claims of the historical argument to be discussed in this chapter, which can be called the *education and burdens argument*, before we assess it:

The Education and Burdens Argument

Moral claim: If an offender had a bad moral education, then there is a reason to impose less hard treatment on him in punishment (than an otherwise similar offender who had a good moral education) *even if he is not less blameworthy for his crime and the right to communicate blame to him held by our courts is not compromised.*

Empirical claim: The proportion of socially deprived offenders who had a bad moral education is greater than that for those who are non-socially deprived offenders.

I will grant the empirical claim of the argument and focus on the moral claim of the argument. Why should bad moral education be relevant to punishment if it does *not* reduce how much censure the offender deserves in our courts? I can see two possible justifications for the moral claim of the education and burdens argument. The first justification appeals to the *burden likelihood principle*: the principle that there is a stronger objection against having a person bear the consequences of his choice if it was *more likely* for him to make the choice, and accordingly, there is a stronger reason to reduce the burden on a person if it was more likely for him to make a choice which invited the burden. If it is true that offenders with a bad moral education are more likely to commit crimes, which are choices that invite the burden of punishment, than offenders with a good moral education, then the burden likelihood principle provides a reason to impose a lesser burden of punishment on an offender with a bad moral education even if he deserves equal blame for his crime from our courts.

The second justification for the moral claim of the education and burdens argument appeals to the *burden opportunity principle*: the principle that there is a stronger reason to reduce the burden on a person if *he had less opportunity to avoid the burden*. If it is true that an offender with a bad moral education had less opportunity to avoid committing his crime and hence incurring the burden of punishment than an otherwise similar offender who had a good moral education, then the burden opportunity principle supports the moral claim of the education and burdens argument.

I will survey and reject the justification based on the burden likelihood principle in section 2 and the justification based on the burden opportunity principle in section 3.

Before I go on, let me briefly compare the two principles. As mentioned in chapter 4, I understand the concept of opportunity in the following manner: How much opportunity a person had to avoid a burden depends on whether (and on how many occasions) he could avoid the burden by making certain choices, how much cost and effort were needed for him to make those choices, and his knowledge of the consequences of his decisions, etc. I will elaborate the concept of likelihood later, in section 2 of this chapter; but even without assuming any particular understanding of the concept of likelihood, the concept of likelihood is clearly distinct from the concept of opportunity. To illustrate, suppose that Ann and Ben have a choice whether to steal. Ann faces much stronger pressure to steal but she is a saint so there is a 100% chance that she will not steal. Ben faces no pressure to steal at all and no one (i.e. 0%) will steal in those circumstances. In such a case, they have unequal opportunity to avoid stealing; but their likelihood of avoiding stealing is the same.

Readers may observe that the two principles that form the main focus of this chapter, namely, the burden likelihood principle and the burden opportunity principle, are *generic* principles that apply to *all sorts of burden including burdens outside criminal punishment*. For example, Roemer accepts something like the burden likelihood principle in the context of distribution of medical costs.[373] That the two principles

[373] Roemer, "Equality and Responsibility".

that form the focus of this chapter can apply to non-punitive burdens should not be surprising. The aim of this chapter is to examine whether there are principles that can justify a reason to reduce the punishment of offenders with a bad moral education *even if they do not deserve less blame*. Given that the special feature of punishment, as compared to other forms of burden, is precisely that punishment involves blame, principles that justify a reason to reduce burden in punishment without justifying the claim that the offender should be blamed less must be principles about *burdens in general, not about burdens in punishment in particular.*

2. Likelihood to make a choice that invited the burden

Some people believe that if a person was more likely to make a choice that invited a certain burden, then there is a stronger objection against him shouldering the burden. For example, illiterate people are more likely to smoke, and smoking increases the chance of cancer greatly; some people believe that provides a reason to give priority to illiterate patients who contracted cancer through smoking than literate patients who also contracted cancer through smoking in distributing medical resources.[374] Analogously, it might be argued that people who had a bad moral education are more

[374] Roemer, "Equality and Responsibility".

likely to commit crimes, a choice which invites the burdens of punishment; and that provides a reason to impose a lesser punishment on an offender who had a bad moral education. The soundness of this argument is the main subject of this section.

Before assessing this argument, I wish to compare this argument with another argument, which is also based on the idea of likelihood and which is sometimes put forward in the literature on moral responsibility and free will. Some people argue that if an offender is more likely to commit a crime, then he is less morally *blameworthy* for his crime. For example, according to this argument, black offenders are less morally blameworthy for their crimes than white offenders who committed similar crimes because black people are more likely to offend. I believe that view on moral blameworthiness is mistaken since whether it is likely for an offender to commit his crime is irrelevant to his quality of will in committing his crime and an offender's moral blameworthiness for his crime depends on his quality of will in doing so. In fact, the idea that likelihood does not affect moral blameworthiness is a very plausible one and is accepted by all free will compatibilists regardless of whether one accepts the quality of will account of moral blameworthiness.[375]

---

[375] See Nelkin, "Difficulty and Degrees of Moral Blameworthiness and Praiseworthiness", 3-5; Morse, "Deprivation and Desert"; Moore, "Causation and the Excuses"; Wolf, *Freedom within Reason*, 103-106; Schoeman, "Statistical Norms and Moral Attributions"; Roemer, *Equality of Opportunity,* 18.

Be that as it may, the issue we are concerned with now is different. To illustrate with a distinction drawn by Scanlon, we have to distinguish attributive responsibility from substantive responsibility.[376] Scanlon believes that a person's attributive responsibility for a choice, i.e. whether it is justified to alter our moral relationship with that person because of his choice, depends simply on his quality of will; but his substantive responsibility for the choice, i.e. whether it is appropriate to ask him to bear the consequences of his choice, depends on other factors.[377] Analogously, even if an offender who is more likely to commit his crime is not less *blameworthy* for his crime, there is a further issue of whether it is less *appropriate to ask him to bear the consequences of committing his crime*; and it is the latter issue that is the focus of this section.

The concept of likelihood or probability can be interpreted in at least three different ways, as suggested by Stephen Perry.[378] The first conception can be called the *relative frequency* conception. According to this conception, likelihood is a property attached to a *reference class of persons* sharing certain characteristics. The likelihood of doing X of a reference class is the relative frequency of doing X for people in that reference class. For example, suppose 35% of white people contracted a disease,

[376] Scanlon, *What We Owe to Each Other*, 248. See also Roemer, *Equality of Opportunity*, 17-18.
[377] Scanlon, "The Significance of Choice".
[378] Perry, "Risk, Harm, and Responsibility", 329.

while only 10% of black people did. In such a case, it is more likely for white people to contract the disease than black people under the relative frequency conception of likelihood. According to the *intrinsic* conception of likelihood, likelihood is linked to the "operation of indeterministic causal processes, such as those associated with quantum mechanics".[379] Suppose, as a matter of quantum physics, an atom has a 60% chance of decaying in the next second; then its intrinsic likelihood of decaying in the next second is 60% regardless of the proportion of similar atoms decaying in the real world. As Perry observes, the intrinsic conception of probability is significant only if we assume the world is *not* causally determined – in a causally determined world the intrinsic probability of every actual event would be 100%.[380] In contrast, we can still sensibly talk about likelihood under the relative frequency conception even if the world is entirely causally determined. The third conception of likelihood can be called the *epistemological* conception. According to the epistemological conception, the likelihood of a person doing X is our *best estimate* of his chance of doing X, *given our existing state of scientific knowledge*. It seems very unlikely that whether it is just to impose a burden on someone would depend on our ability to predict, and accordingly how likely he would incur the burden in the epistemological sense. Moreover, there does not seem to be much reason to believe that offenders with a bad moral education

[379] Perry, "Risk, Harm, and Responsibility", 323.
[380] Perry, "Risk, Harm, and Responsibility", 324.

have a higher intrinsic likelihood of offending.[381] So I will focus on the relative frequency conception of likelihood in the rest of this section.

Is likelihood in the relative frequency sense morally significant? I believe that relative frequency is *clearly* morally insignificant unless we describe the characteristics that define the reference classes in the most detailed manner.[382] Let me illustrate this point. Suppose Asians have a higher relative frequency to smoke, a choice that invites lung cancer, than non-Asians. Arthur is Asian and Betty is not. They both contracted lung cancer through smoking. Do we have a stronger reason to relieve Arthur of the burden – say, if we only have one pill, should we give it to Arthur? I think the answer is clearly "not necessarily", *even if* we assume that a person's likelihood, under the relative frequency conception, of making a choice, *in some sense,* is relevant to whether he should bear the consequences of the choice. This can be seen clearly if we add that while Asians have a higher relative frequency of smoking than non-Asians, *Asian men* have a lower relative frequency of smoking than non-Asian women. In such a case, *even if* we assume that a person's likelihood, under the relative frequency conception, of making a choice, *in some sense*, is relevant to whether he should bear

[381] Buss, "Justified Wrongdoing", 339. Moore, "Causation and the Excuses". For a reply that I do not regard to be successful, see Kaye, "Resurrecting the Causal Theory of the Excuses"; Kaye, "The Secret Politics of the Compatibilist Criminal Law", 403-405.
[382] Perry, "Risk, Harm, and Responsibility", 333-336.

the consequences of the choice, it seems wrong to focus on the reference class of "Asians" rather than "Asian men" and give priority to Arthur on the ground that he is an Asian. But of course, it is arbitrary to stop at "Asian men" – why not "Asian young men"? Logic seems to dictate that we have to look for the reference class defined in the *most detailed manner.*

What, then, is the reference class defined in the most detailed manner that a person belongs to? It must be a group that contains people who share *all* the characteristics, internal and external, temporal and territorial, of that person. Since no two persons share all characteristics, a reference class that is defined in the most detailed manner must be a single-person group. Accordingly, for every individual who actually made a choice (say, the choice to smoke), the relative frequency to make that choice for the reference class that is defined in the most detailed manner that he belongs to is the same, namely, 100%. As such, the relative frequency conception of likelihood cannot be used to justify the relevance of moral education to the distribution of punishment – since no matter whether an offender had a good moral education or not, his likelihood of offending, a choice which invited the burden of punishment, under the relative frequency conception of likelihood and taking the reference class that is defined in the most detailed manner, would be 100%.

My claim that a person's likelihood to make a certain choice, under the relative frequency conception of likelihood, is morally irrelevant to whether it is fair to ask him to bear the consequences of his choice, may sound surprising to some. So let me consider a rejoinder here based on the work of Roemer who famously links the issue of substantive responsibility with the idea of likelihood under the relative frequency conception. Roemer sometimes suggests that whether and how much a person should bear the consequences of a choice (like the decision to smoke) depends on what proportion of people in the same reference class would make the same choice; if most people in that reference class would make the same choice, then he should bear fewer consequences of the choice.[383] Reference classes, according to Roemer, are defined in terms of *a finite list* of characteristics,[384] including race, housing location, education level, parental education level, etc.[385]

To illustrate, suppose that two persons, Bill and Lily, both chose to smoke; and as a result both contracted lung cancer. We only have medical resources to help one of them. Bill is a black person who lived in a poor neighbourhood. He dropped out of high school and his parents did not graduate from high school either. Lily is a white

[383] Roemer, "Equality and Responsibility".
[384] Roemer, "Defending Equality of Opportunity", 263.
[385] Roemer, *Equality of Opportunity*, 13.

college graduate who lived in a suburban house. Both of her parents graduated from college. In deciding who should be given priority, Roemer would ask what proportion of black persons who lived in a poor neighbourhood and who dropped out of high school and whose parents did not receive a college education would make the choice of smoking and what proportion of white university graduates who lived in a middle class neighbourhood with parents who graduated from college would make the same choice. If the percentage is higher for the former group (which I assume to be the case), then Roemer claims there is a stronger objection against letting Bill bear the consequences of his choice and hence we should give priority to Bill.[386]

For our purpose here, the most significant thing about Roemer's proposal is that while he narrows down the reference classes in a certain way (e.g. he does not simply compare black persons to white persons but instead takes into account a list of other characteristics like education level), he does *not* narrow the reference classes down to the narrowest possible groups (i.e. single-person reference classes).[387] I think most people agree that Bill should be given priority in this case. If that is true, then doesn't it support the idea that a person's likelihood of making a choice, under the relative frequency conception of likelihood, is relevant to whether it is fair to ask him to bear

[386] Roemer, "Equality and Responsibility".
[387] Roemer, "Defending Equality of Opportunity", 6.

the consequences of his choice, *even if we do not define reference classes in the narrowest possible manner*?

Let me briefly state my reply before elaborating it. I believe that the reason why a person's likelihood to make a choice, under the relative frequency conception of likelihood and given the way in which Roemer defines reference classes, is *in many cases* relevant to whether it is fair to ask him to bear the consequences of his choice is because such likelihood *negatively correlates* with a property that many people believe to be morally significant, namely, *opportunity* to avoid incurring the burden. But likelihood is morally insignificant when it does not affect opportunity and this is clearest in the case where the two ideas come apart.

Let me now elaborate my reply. What is the rationale for defining the reference classes in Roemer's manner? Why does he take some properties into account (race, social class, education level, etc.) but not others (date of birth, date of making the choice, location of making the choice, etc.)? The obvious answer is that the former set of properties affects the agent's *opportunity* to avoid the burden.[388] People who live in a poor neighbourhood usually face stronger peer pressure to smoke, and given that

[388] This idea seems congenial to Roemer, given that his theory, after all, is offered as an account of equality of "opportunity".

a person’s opportunity to avoid a burden is affected by how difficult it was for him to make the choice (or choices) that would enable him to avoid the burden, they have less opportunity to avoid incurring the burden of cancer; but there is no obvious reason why people who are born in December would face stronger pressure to smoke than people who are born in October.

In many cases, the assumption that people who have more opportunity to avoid incurring a burden are less likely to make a choice that invites the burden (and vice versa) will hold. For example, people who faced stronger pressure to smoke (and therefore had less opportunity to avoid cancer) are more likely to smoke. When this assumption holds, the reference classes (as defined by Roemer) that have a higher relative frequency of making a choice which invites a burden *will happen to be* the reference classes that consist of people who had less opportunity to avoid incurring the burden. Since in many cases a person’s opportunity to avoid incurring a burden is relevant to whether it is fair to ask him to shoulder the burden (but see section 3 of this chapter below), whether a person falls under a reference class (as defined by Roemer) that has a higher relative frequency of making a choice that invites a burden is, in many cases, relevant to whether it is fair to ask him to shoulder the burden.

Yet, the assumption that likelihood is negatively correlated to opportunity to avoid does not always hold. And when the two come apart, it is clear that likelihood does not matter. To see this, suppose that despite the fact that black people who come from poor families with uneducated parents face stronger pressure to smoke, they have a lower relative frequency of smoking and thus contracting cancer. (Perhaps purely as a matter of coincidence, they are the group of people who are most careful about their health.) If likelihood under the relative frequency conception is what matters, then we should give priority to people in privileged background in distributing medical resources. But surely that is wrong, as suggested by Risse, Mason, and Maskin.[389] In fact, Roemer seems to accept this point when he concedes that statistical likelihood, under the relative frequency conception, is only a *proxy* to measure opportunity.[390]

(Let me clarify what I am arguing for here. I am *not* arguing that opportunity *always* matters, i.e. there is *always* an extra objection against letting a burden fall on someone who had less opportunity to avoid incurring it. In fact, I will argue against that claim in section 3 below. My point here is simply that there is no extra objection against asking a person who had a higher likelihood (in Roemer's sense of likelihood) of making a burden-inviting choice to bear the consequences of his choice, *if he does not*

---

389 Risse, "What Equality of Opportunity Could Not be", 741-742; Mason, "Equality, Personal Responsibility, and Gender Socialisation", 237; Maskin, "Reply to Roemer on Equality".
390 Roemer, "John Roemer Replies".

happen to be someone who had less opportunity to avoid incurring the burden. If what I argue in section 3 below is correct, then *even if* a person who had a higher likelihood (in Roemer's sense of likelihood) of making a burden-inviting choice happens to be someone who had less opportunity to avoid incurring that burden, there may still be no extra objection against asking him to bear the consequences of his choice.)

3. Opportunity and burden

The second justification for claiming that there is a reason to reduce the burden of punishment on offenders who had a bad moral education appeals to the concept of *opportunity* rather than the concept of likelihood. As an analogy, let us start by considering the smoking case introduced at the beginning of section 2 of this chapter again. It might be suggested that illiterate smokers should indeed bear lower medical costs in case they contract lung cancer as compared to literate smokers; but the reason why illiterate smokers should bear lower cost for their lung cancer is not because it is more likely for illiterate people to smoke and thus to contract cancer; rather it is because, regardless of likelihood, they had less opportunity to avoid smoking and thus to avoid contracting cancer. Similarly, one may argue, as Litton does, that offenders with a bad moral education had less opportunity to avoid committing crimes and thus incurring the burdens of punishment; and that provides a reason to reduce their burdens in punishment, regardless of whether it was more likely for them to offend.[391] Rejecting this argument for claiming that there is a reason to reduce punishment of offenders with a bad moral education is the task of this section.

As mentioned above in chapter 4, I understand the concept of opportunity in the following manner: How much opportunity a person had to avoid a burden depends on whether (and on how many occasions) he could avoid the burden by making certain choices, how much cost and effort were needed for him to make those choices, and his knowledge of the consequences of his decisions, etc. Let me make one point about the *historical* aspect of the concept of opportunity (which I also touched on in chapter 4 above). In this chapter, I am examining whether moral education has any *intrinsic* significance to punishment; so I am comparing offenders with a good moral education and offenders with a bad moral education who were nonetheless identical *at the time of offence*. Insofar as their *present opportunity* to avoid committing their crimes (i.e. their opportunity *at the time of offence* to avoid committing their crimes), they must have the *same* amount. Yet, that does not mean that we cannot sensibly say that they may have a different amount of opportunity to avoid committing their crimes because

[391] Paul Litton, in arguing for the relevance of moral education to punishment, focuses on opportunity but not likelihood: Litton, "The 'Abuse Excuse' in Capital Sentencing Trials", 1065. See also Nussbaum, "Equity and Mercy".

one can have, *in addition* to present opportunity to avoid, *past* opportunity.

To illustrate the concept of *past* opportunity, consider two drivers who both needed to apply the brake in order to avoid a car crash. The first driver was unconscious at t1; he woke up at t2; if he did not apply the brake at t2 the car would immediately crash. The second driver was identical except that he was conscious at t1. If he applied the brake at *either* t1 *or* t2, the car crash would be avoided. Here the second driver had more opportunity to avoid the car crash even though the two drivers had equal present opportunity to avoid the car crash (they were identical at t2 when the car crash happened) because the second driver had more *past* opportunity to avoid the car crash. As another example to illustrate the historical component of opportunity, a student who received a good academic education may have more opportunity to acquire academic knowledge, and thus more opportunity to avoid doing badly in an exam, even if his constitution and situation were identical to another student at the time of exam.

It seems to me plausible that offenders with a bad moral education had less *past* opportunity to avoid committing their crimes, and accordingly, to avoid incurring the burden of punishment. For example, it seems plausible that the son of the Mafia boss faced much stronger pressure to associate with criminals when he was young; and of course associating with the wrong kind of people when one was young is criminogenic. In any case, I will assume that offenders with a bad moral education had less past opportunity to avoid committing their crimes for the sake of argument. The question I want to answer is whether it is true that if a person had less opportunity to avoid incurring a burden, then there *must* be a stronger objection against letting the burden fall on him. I believe the answer is no.

It is undeniable that the amount of opportunity one had to avoid a burden is *sometimes* relevant to whether it is just to ask him to shoulder the burden. Imagine two otherwise identical patients. One of them contracted a disease because a mad scientist forcefully injected the virus into his body. The second patient, in contrast, injected the virus into his own body voluntarily because he thought it was fun to do so. In this case, the first patient, unlike the second patient, did not have any opportunity to avoid incurring the burden, i.e. the disease. It seems plausible that if we only have one medicine we should give it to the first patient.

However, I do *not* believe that we should generalise from that case and conclude that there is *always* a stronger objection against letting an individual to shoulder a burden if he had less opportunity to avoid the burden. I believe that there are four main justifications for why *sometimes* we should make falling of burdens sensitive to people's opportunity to avoid them; but those four justifications do *not* show that there is *always* a stronger objection against letting an individual to shoulder a burden if he had less opportunity to avoid that burden. Moreover, most relevant for our purpose here, I believe that none of those four justifications suggests that there is a stronger objection against imposing the burden of punishment on an offender just because he had less past opportunity to avoid committing his crime and thus incurring the burden of punishment.

Let me elaborate my position, starting by asking *why sometimes* we should make falling of burdens sensitive to people's opportunity to avoid them. The first has to do with the idea of *consent, or the wider notion of forfeiture*.[392] This idea can account for the injection case above: since the second patient voluntarily took the risk of contracting the disease with full knowledge of the consequences, he has forfeited some of his rights against the harm; accordingly, he has fewer complaints about contracting the disease. Since in general a person who had no opportunity to avoid a burden at all cannot be said to have voluntarily taken the risk of incurring the burden,

[392] Scanlon, *What We Owe to Each Other,* 258.

the amount of opportunity one had to avoid incurring a burden is sometimes relevant to whether there is an extra objection against asking him to shoulder the burden.

The idea of forfeiture, however, may not account for the full moral importance of the idea of opportunity. Imagine two persons who are *otherwise identical*. Both of them can avoid a disease by taking a certain vaccine. The vaccine would cause *minor* discomfort to the first person, while the vaccine would cause moderate but still not serious discomfort to the second person. Suppose that both of them did *not* take the vaccine, and as a result, contracted the disease. (Call this case the "original vaccine" case.) Many people believe that if we only have one medicine, there is a stronger reason to give it to the second person.

This belief cannot be accounted for by the idea of forfeiture. This is because forfeiture seems to involve what Rawls referred to as a "range property":[393] in general,[394] one cannot validly forfeit his rights against a burden without any opportunity to avoid the burden *but* the validity and normative force of one's forfeiture are *not directly proportional* to the amount of opportunity to avoid the burden he had. As an analogy, a store owner's consent to sell a book is not less valid than another store owner just

[393] Rawls, *A Theory of Justice: revised edition*, 444.
[394] I set aside exceptional cases like the Frankfurt cases.

because his bank account has slightly less money so he faces more economic pressure to sell the book. It seems that, given that the cost to take the vaccine is not that onerous, both persons in the original vaccine case can be said to have taken the risk to contract the disease voluntarily and thus their positions would be the same *insofar as forfeiture is concerned*.

What then can explain the priority to the second person in the original vaccine case? Let me first set aside two suggestions that do not work. First, it may be suggested that the first person is more *culpable* in taking the risk. But this suggestion does not work as I do not think that the first person is more culpable. It need not be culpable for a person to avoid discomfort even if it is really minor.

Second, it may be suggested that *if it would be more costly* for a person to avoid a burden, *then there is always* a stronger reason against letting the burden fall on him. But that does not seem to be true in general. We can see this by considering *a variation* on the original vaccine case, which can be called the "amended vaccine" case. *Instead of* imagining that the two patients are identical except their level of discomfort in taking the vaccine, suppose that the first person (but not the second person) was born with cancer (which is unrelated to the disease preventable by the vaccine) and has suffered serious discomfort throughout his life because of cancer. I have a *different* view about the amended vaccine case as compared to the original vaccine case. It seems to me that given the discomfort that the first person has suffered in the past, it is less fair to ask him to bear the extra discomfort of taking the vaccine, even though the extra discomfort of taking the vaccine is less serious to him as compared to the second person; and as a result it is less fair to ask the first person to bear the burden of the disease as compared to the second person in the amended vaccine case. This intuition seems to be shared by others. Tadros, for example, agrees with me that in deciding whether it is fair to ask a person to bear the consequences of a certain choice, we need to look at whether it is fair to ask him to choose otherwise, and whether it is fair to ask him to choose otherwise depends, *in addition to his cost of choosing otherwise in that particular case, on more global considerations like how much burden he has shouldered in the past*.[395]

If it is neither forfeiture, nor culpability, nor the simple idea that it is *always* less fair to ask a person to shoulder a burden if it would be more costly for him to avoid it, then what can explain why priority should be given to the second patient in the original vaccine case (i.e. the case where none of the two patients has cancer and both

[395] Tadros, *The Ends of Harm*, 174.

of them are identical except that the vaccine would cause a greater discomfort to the second patient)?

Perhaps we can draw help from Scanlon's work. Scanlon's famous "value of choice account" appeals to the idea that an opportunity (or a better opportunity) to avoid a burden is in itself valuable to an agent (i.e. an opportunity is valuable even if it is not exercised), rather than the idea of forfeiture, in explaining why falling of burdens should be sensitive to people's opportunity to avoid them.[396] The value of choice account seems to be based on *reciprocal fairness*, the idea that we should match benefits with burdens: if a person received more benefits then there is a weaker objection against him shouldering burdens. Since, in the *original* vaccine case, the first patient had a better opportunity to avoid the burden, which is a form of benefit in itself, and he is otherwise identical to the second patient, the first patient has received more benefits; and accordingly he has a weaker objection against bearing the cost of the disease. The idea of reciprocal fairness also explains why, in contrast, we should give priority to the first person in the *amended* vaccine case: given his cancer, he did not receive more benefits in total.

[396] Scanlon, *Moral Dimensions*, 205-206; *What We Owe to Each Other*, 251-256; "The Significance of Choice"; "Punishment and the Rule of Law", 230. But for objections to the idea that having more opportunity is *per se* beneficial, see Williams, "Liberty, Liability, and Contractualism", 284; Voorhoeve, "Scanlon on Substantive Responsibility".

The ideas of forfeiture and reciprocal fairness account for many cases where we want to make falling of burdens sensitive to people's opportunity to avoid them. For example, I believe that these two ideas explain why in many cases we want to give priority to illiterate smokers over literate smokers in distributing medical costs: illiterate people are usually those who received least benefits in society in total so reciprocal fairness would suggest giving more medical resources to them (at least if that does not involve harming other worse off people); some of the illiterate smokers may be compulsive smokers who cannot quit smoking voluntarily after their teenage years (so they did not make any voluntary right-forfeiting choices in relation to smoking when they were adults) and, given their family influence, they picked up smoking when they were too young to make any right-forfeiting choices.[397] Yet, the two ideas, forfeiture and reciprocal fairness, *are not the full story* in explaining why falling of burdens should sometimes be sensitive to people's opportunity to avoid them. We can see why these two ideas are not the full story by thinking about the following case. Suppose we can choose between two systems. First, we allow people to create liabilities on themselves by creating legally recognized contracts, as long as

[397] On the other hand, if the illiterate smoker is rich and he was not a compulsive smoker when he was an adult so he could quit smoking voluntarily any time when he was an adult, the fact that he faced stronger pressure to smoke (and accordingly had less opportunity to avoid smoking) would not provide a reason to give him priority in distributing medical resources. See my discussion on punishment below.

they were voluntary in making the contracts. Second, we do not allow people to do so by not recognizing contracts in our law.

The first thing to observe is that the falling of burdens is more sensitive to people's opportunity to avoid them under the first system. While under both systems people who *lack* the opportunity to avoid contractual liability do *not* incur any contractual liability, only under the first system can people who *have* the opportunity to avoid contractual liability incur contractual liability by his choices.

But the next thing to observe is that if we value freedom of contract at least sometimes, and hence prefer the first system over the second at least in some cases, our preference *cannot* be explained on the ground of forfeiture or reciprocal fairness. In neither system would contractual liability fall on someone who had not voluntarily accepted the risk of incurring contractual liability, so the idea of forfeiture provides no basis for preferring the first system over the second.[398] Moreover, a system that provides for freedom of contract need not promote reciprocal fairness, as Cohen observed.[399] So this case suggests that there are grounds to make falling of burdens sensitive to people's opportunity to avoid them, over and above the ideas of forfeiture

398 For a similar point, that forfeiture does not provide a reason to choose between compulsory insurance and non-compulsory insurance, see Barry, "Chance, Choice, and Justice", 148.
399 See Cohen, *Self-Ownership, Freedom, and Equality.*

and reciprocal fairness.

What then can explain why we prefer some freedom of contract over no freedom of contract? One immediate answer is that we want to give people more legal power, but that answer is not satisfying: *why* should we give people more legal power? I do not think there is a general moral principle that requires us to give as much legal power to people as possible. Rather, I think, sometimes we should give people more legal power because of two reasons. First, doing so may promote *utility.* Second, more importantly, doing so may promote *autonomy*.[400] The idea of autonomy explains why we want to provide people with the legal power to make at least some legally enforceable contracts. Just imagine how a diversity of ways of living, which presupposes a division of labour, can be possible without any contracts recognized by and enforceable in law.

We have discovered four grounds for why *sometimes* falling of burdens should be sensitive to people's opportunity to avoid them: forfeiture, reciprocal fairness, utility, and autonomy. The question now is whether these four grounds suggest that there is always an extra objection (even if a defeasible one) against imposing the burden of

[400] See, e.g. Dagan, "Autonomy, Pluralism, and Contract Law Theory". For a review of theories of contract that appeal to the concept of autonomy, see Benson, "Contract", 37.

punishment on someone who had less past opportunity to avoid committing his crime and hence incurring the burden of punishment.

I think the answer is no. The idea of forfeiture cannot justify the claim that there is always a stronger objection against imposing the burden of punishment on an offender who had less past opportunity to avoid committing his crime and hence incurring the burden of punishment. This is because, as mentioned above, forfeiture is a range property. An offender who, at the time of offence, was not acting under factual/legal ignorance, duress, or other forms of compulsion that negate voluntariness, because of the amount of present opportunity he had, generally had *sufficient* opportunity to avoid committing his crime and hence incurring the burden of punishment for the purpose of forfeiture *even if* he did not have any past opportunity to avoid committing his crime.

An analogy may clarify my point here. Suppose I had two opportunities to avoid a certain contractual liability to a company (say, the company called me twice, and if I cancelled the contract in any one of the two occasions then I would be relieved from contractual liability) and you only had one opportunity (say, the company only called you once). Suppose we were not under duress, compulsion, etc., when we were called, and we both did not cancel the contract. It seems that the company has an *equal* right against us since we *both voluntarily* accepted the contractual liability, *despite the fact that I had more past opportunity to avoid the contractual liability.* Analogously, even if the offender who had a bad moral education had less past opportunity to avoid committing his crime, given that he had sufficient opportunity to avoid the relevant criminal liability at the time of offence (by not offending), he can be said to have taken the risk of being punished voluntarily; accordingly the idea of forfeiture would not justify drawing a distinction between him and an otherwise similar offender who had a good moral education. (The analysis would, of course, be more complicated for offenders who acted under factual/legal ignorance, duress, or compulsion that negates voluntariness at the time of offence since they cannot be said to have voluntarily taken the risk of being punished when they offended – but clearly they are not the central cases of offenders.)

I wish to consider one objection here. I have said that forfeiture is a range property; *for the purpose of forfeiture,* once a person's opportunity to avoid a burden was sufficient, i.e. his amount of opportunity was above a certain threshold, how much extra opportunity he had does not matter. I then said that an offender who was not acting under ignorance of law/fact and compulsion, etc. had sufficient opportunity to avoid incurring the burdens of punishment, for the purpose of forfeiture. My opponent may object, however, that even if we accept that forfeiture is a range property, I have not told the readers where exactly the threshold lies and why we should set the threshold there. As such, how do I know that as long as an offender was not acting under ignorance of law/fact and compulsion, etc., then he had sufficient opportunity to avoid the burdens of punishment, for the purpose of forfeiture?

In reply, I accept that I do not have a theory for setting the appropriate threshold at present. I hope to be able to provide that on another occasion. Nonetheless, I have two reasons for believing that, no matter where exactly the threshold ought to be set, an offender who was not acting under ignorance of law/fact and compulsion, etc. had sufficient opportunity to avoid the burdens of punishment for the purpose of forfeiture.

The first is simply that I see no rational distinction between him and a gambler in a casino. Gamblers may face stronger or weaker temptations and pressures, act under different degrees of uncertainty, and may have different histories; but if they are not acting under ignorance, serious threats, or involuntariness, I think most people will regard them as having sufficient opportunity to avoid losing and would hold them responsible for their bets.

The second point is that the proper level of threshold may, as Feinberg pointed out, be context-sensitive – it depends on what burden we are talking about.[401] And there are some considerations to suggest that we should not insist on a very high threshold in the context of a liberal system of punishment. First, a liberal polity respects persons as choosing agents; and this is inconsistent with setting a very high threshold as that would imply that people should not be held responsible for most of their choices.[402] Second, none of the major purposes of criminal punishment, like retribution or deterrence, suggests that we should set a very high threshold.

Litton focuses on reciprocal fairness in arguing that there is a reason for an offender who had less past opportunity to avoid committing his crime, and thus incurring the burdens of punishment, to receive less punishment even though he is not less blameworthy for his crime.[403] But, in reply, reciprocal fairness cannot justify the claim that there is always a stronger objection against imposing the burdens of punishment on someone who had less past opportunity to avoid them because opportunity to avoid incurring the burdens of punishment, even assuming that it is a

[401] Feinberg, *Harm to Self*, chapter 20.
[402] Bennett, "The Limits of Mercy"; Duff, "The Intrusion of Mercy", 375-376.
[403] Litton, "The 'Abuse Excuse' in Capital Sentencing Trials", 1066.

kind of benefit, is only one kind of benefits amongst many; there is no reason to believe that a person who had less past opportunity to avoid committing crimes and hence incurring the burdens of punishment has less goods in total (recall the son of the rich Mafia boss) and hence deserves less burden as a matter of reciprocal fairness.

Let me consider a rejoinder by my opponent here. My opponent may suggest that I wrongly assumed that reciprocal fairness *must* take a *holistic* form, looking at *all the benefits and burdens a person has in his life in total* in deciding whether it is justified to ask him to shoulder a further burden. My opponent may suggest that, regardless of whether there is a holistic form of reciprocal fairness, there is a *localized form of reciprocal fairness,* which aims to *match benefits and burdens related to a particular practice.*[404] Since a person who had less opportunity to avoid offending had fewer "safeguards" against incurring the burden of punishment,[405] which are a form of benefits *related to the practice of criminal punishment*, there is a stronger reason to reduce the amount of burdens imposed on him in criminal punishment *no matter how many benefits and burdens he received in other spheres of his life.*

[404] The idea that there is a form of localized reciprocal fairness as applied to punishment may be implicit in Litton, "The 'Abuse Excuse' in Capital Sentencing Trials", 1066. The idea of localized reciprocal fairness in punishment has been discussed in Bayern, "The Significance of Private Burdens and Lost Benefits for a Fair-play Theory of Punishment", 6-10. See also Temkin, "Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck", 65-66.

[405] Litton, "The 'Abuse Excuse' in Capital Sentencing Trials".

I have two replies to this rejoinder by my opponent, which appeals to the concept of *localized reciprocal fairness*. First, *even if* we assume that there is a localized form of reciprocal fairness, i.e. there is a *pro tanto* reason to match burdens *related to a particular practice* with benefits *related to that particular practice*, it does not follow that there is always a stronger objection against imposing the burdens of punishment on offenders who had less opportunity to avoid offending. This is because there are *many* benefits and burdens related to the practice of criminal punishment other than safeguards against incurring the burdens of punishment. For example, the son of the rich Mafia boss may have received fewer safeguards against incurring the burdens of punishment; but he may also have received extra benefits from the system of criminal punishment because he has more properties, which are protected by the existence of the system of criminal punishment, such that he need not be the one who has received fewer goods in total *even if we count exclusively benefits and burdens related to the practice of criminal punishment*.

Second, and more importantly, I do not think that localized reciprocal fairness is a valid moral principle. In other words, I reject the idea that there is a reason to match benefits and burdens related to a particular practice. The poor people in our society benefit more from a public healthcare system than the rich people – does it follow that there is a reason to let the poor people shoulder the cost of such a system? I do not think, given the global distribution of benefits and burdens, the rich people have *any* legitimate complaint at all if we put the cost of setting up the public healthcare system on them.

I suspect the people who accept the principle of localized reciprocal fairness do so because they confuse the principle of localized reciprocal fairness with some morally valid principles that look similar to it but which are in fact different. Let me explain. I *accept* that there are valid *localized fairness* principles.[406] By localized fairness principles I mean principles of fairness *that do not depend on the wider social distribution of benefits and burdens*. For example, if you jump the queue in a bus stop, then it is unfair in some sense to other people who are queueing *regardless of* the wider social distribution of benefits of burdens.[407] And if we punish an innocent person, then that is unfair in some sense no matter what the wider social distribution of benefits and burdens is. But these localized fairness principles are *not* principles about *reciprocal* fairness, i.e. they are *not* principles about *matching burdens with benefits.* Rather, they are about fairness in two other senses: that we should follow established conventions within a practice and that we should pay attention to the

[406] See section 2.2.1 of chapter 8 of this thesis above.

[407] Bayern, "The Significance of Private Burdens and Lost Benefits for a Fair-play Theory of Punishment", 8.

meaning of the goods and burdens within a practice in distributing them. This can be seen clearly once we pay a closer attention to the fairness norms in the bus queue case and the punishment case: the unfairness of jumping the queue does not depend on whether the queue jumper is a slow walker such that he is at the end of the queue every time and thus receives less benefit *related to the practice of queuing*; nor would the unfairness of punishing the innocent be mitigated if we find out that the innocent person lives in a neighbourhood that is very well-protected by the police such that he benefits more from the existence of the system of criminal punishment than other citizens.

How about utility and autonomy? While, as we have seen above, in some cases by making falling of burdens sensitive to people's opportunity to avoid them we can promote utility and autonomy, it is very unclear how making burdens of punishment sensitive to people's past opportunity to avoid committing crimes, and thus incurring the burdens of punishment, can promote autonomy and utility. Doing so, unlike giving people the power to make legally valid contracts, would not enable some extra valuable ways of living.

4. Conclusion

In this chapter, I have criticized the education and burdens argument. According to this argument, there is a reason to punish an offender who had a bad moral education less for his crime (than an otherwise similar offender who had a good moral education) *even if* he is not less blameworthy for his crime and the right to communicate blame to him held by our courts is not compromised; and the proportion of socially deprived offenders who had a bad moral education is greater than that for those who are non-socially deprived offenders.

I have focused on attacking the moral claim of the argument. I have explored two possible justifications for the claim. The first appeals to the concept of likelihood: it is more likely for people with a bad moral education to commit crimes, and there is a stronger reason to reduce a burden on a person if it was more likely for him to make a choice that invited the burden. The second appeals to the concept of opportunity: an offender with a bad moral education had less opportunity to avoid committing his crime and hence incurring the burdens of punishment, and there is a stronger reason to reduce a burden on a person if he had less opportunity to avoid incurring that burden. I have argued that both justifications are unsound.

Chapter 10: Whole life moral desert and payment in advance

1. Introduction

Now we have come to the last of a trio of chapters (chapters 8 to 10) constituting part 4 of this thesis. To recall, part 4 is concerned with non-exculpatory non-communicative arguments, i.e. arguments for claiming that there is a ground to punish an offender less for his crime, that is based neither on the idea that the offender is less blameworthy for his crime nor the idea that our courts' standing to communicate blame to him is compromised, that is more applicable to socially deprived offenders.

In chapter 8, I have examined a non-exculpatory non-communicative argument that appeals to the idea of distributive justice (in the sense of improving the well-being of the worst off members of society); in chapter 9, I have examined another non-exculpatory non-communicative argument that appeals to the idea of moral education. In this chapter, I will explore two more non-exculpatory non-communicative arguments.

According to the first argument, which can be called the *whole life moral desert argument* and which I constructed based on the work of Ezorsky, there is a reason to promote *whole life moral desert*, i.e. the state of affairs where the total level of misery in each person's whole life corresponds to the total level of moral blameworthiness displayed in his whole life.[408] Given that socially deprived offenders have already suffered much in their lives, by punishing a socially deprived offender less we are likely to promote whole life moral desert while that is not the case if we reduce the punishment of a non-socially deprived offender.

According to the second argument, which can be called the *payment in advance argument* and offered by Klein, there is a reason to reduce the penalty imposed on an agent for an act if the mental states that motivated the act were acquired as a result of a more painful process.[409] Since generally speaking the upbringings of socially deprived offenders, which led to the acquisition of mental states that motivated their crimes, are more painful, socially deprived offenders are likely to deserve a less severe penalty for their crimes.

---

[408] Ezorsky, "The Ethics of Punishment", xxii-xxvii. Ezorsky accepts the moral claim of the argument. But it is unclear whether she accepts the empirical claim of the argument.
[409] Klein, *Determinism, Blameworthiness, and Deprivation*.

I want to briefly explore the relationship between the two arguments to be discussed in this chapter and the other two arguments discussed in part 4 of this thesis. We can distinguish between two kinds of non-exculpatory non-communicative argument. The first kind is *consequentialist*: this kind of argument claims that by reducing the punishment of a socially deprived offender we are more likely to *cause a certain good outcome* (like the improvement of the well-being of the worst off members of society) than by reducing the punishment of a non-socially deprived offender. The second kind is *non-consequentialist*: this kind of argument claims that if an offender possesses a certain feature (or if he possesses that feature to a higher degree) then there is a reason to punish him less for his crime *even if doing so will not cause any good consequences*, and the proportion of socially deprived offenders who possess that feature is greater than that for those who are non-socially deprived offenders (or that on average socially deprived offenders possess that feature to a higher degree than non-socially deprived offenders).

The improvement of the worst off argument, as I have observed in chapter 8, is a consequentialist argument. The education and burdens argument, in contrast, is a non-consequentialist argument. The whole life moral desert argument, like the improvement of the worst off argument, is a consequentialist argument. The payment in advance argument, like the education and burdens argument, is a non-consequentialist argument. In the previous two chapters, I have rejected the education and burdens argument and I have stated my reservations about certain aspects of the improvement of the worst off argument.

Given that the whole life moral desert argument falls into the same category as the improvement of worst off argument and the payment in advance argument falls into the same category as the education and burdens argument, an obvious question to ask is whether they can avoid the difficulties faced by the two arguments that I discussed in the previous two chapters. I will show, in the remaining parts of the chapter, that the answer is no. In particular, I believe that the whole life moral desert argument, in addition to being vulnerable to the same difficulties faced by the improvement of the worst off argument, is vulnerable to *additional* problems *not* faced by the improvement of the worst off argument.

2. The whole life moral desert argument

Some people agree with Ross and Ezorsky that whole life moral desert is a good state of affairs, i.e. it is good where the total level of misery in each person's whole life corresponds to the total level of moral blameworthiness displayed in his whole life.[410] The idea of whole life moral desert may seem to be able to explain why there is often a reason to punish a socially deprived offender less for his crime. We should start by observing that the idea of whole life moral desert does *not* distinguish between suffering inflicted by criminal punishment and *other suffering*. It is a good thing, according to this idea, if a brutal serial rapist who happily escaped his legal punishment and who would otherwise live a wonderful life got a car accident, which inflicted suffering on him. Since, according to this argument, socially deprived offenders have already suffered much in their lives, by reducing their punishment we promote whole life moral desert.[411] In contrast, since non-socially deprived offenders have not suffered much in their lives, the idea of whole life moral desert does not call for a reduction of their punishment. Let me lay out this argument in a schematic manner before we assess it.

<u>The Whole Life Moral Desert Argument</u>

Moral claim: If by reducing the punishment of an offender we will promote whole life moral desert, then there is a reason to reduce his punishment that does not apply to offenders the reduction of punishment of whom will not lead to the promotion of

[410] Ross, "Punishment", 147; Ezorsky, "The Ethics of Punishment", xxiv.
[411] Ezorsky, "The Ethics of Punishment", xxii-xxvii. See also Card, "On Mercy", 182; Goldman, "Toward a New Theory of Punishment", 61.

whole life moral desert; and accordingly there is a reason to punish him less for his crime than an otherwise similar offender the reduction of punishment of whom will not lead to that consequence.

Empirical claim: For a greater proportion of socially deprived offenders than non-socially deprived offenders, reducing their punishment will promote whole life moral desert.

The whole life moral desert argument, as I said above, is a consequentialist argument like the improvement of the worst off argument. However, they are not identical as they appeal to *different good consequences.* For the improvement of the worst off argument, the good consequence that it tries to promote by mitigation of punishment is *distributive justice* whereas the good consequence that the whole life moral desert argument tries to promote by mitigation of punishment is *whole life moral desert*. The two states of affairs are different for at least two reasons. First, it seems plausible that distributive justice should, to a certain degree, be sensitive to option luck while that is not true for whole life moral desert.[412] Second, if a person voluntarily sacrifices himself for others with full knowledge of the consequences, from the point of view of

[412] Dworkin, *Sovereign Virtue*, chapter 2. But c.f. Cohen, "Fairness and Legitimacy in Justice".

distributive justice he may be entitled to less but from the point of view of whole life moral desert he does not deserve less.[413]

Now let me recall my reservations about the improvement of the worst off argument, which I raised in chapter 8 above. I accepted the moral claim of the improvement of the worst off argument. I also acknowledged that the empirical claim of the argument has some intuitive appeal: it may be thought that given that socially deprived offenders are the worst off members of society and non-socially deprived offenders are not, mitigating the punishment of socially deprived offenders, unlike mitigating the punishment of non-socially deprived offenders, will benefit the worst off members of society.

But I have also argued, in chapter 8 above, that chain of reasoning is not solid once we take into account the *indirect consequences* of mitigation of punishment. First, for *some* socially deprived offenders, punishing them less will not improve their own well-being. Second, for *some* socially deprived offenders, punishing them less will increase crimes, and hence, given that most victims of crimes are also worst off

[413] It is of course possible for a theory of distributive justice to take into account moral desert, though it is not usual. For an outline of such theories, see Arneson, "Luck Egalitarianism – A Primer", 146. See also Stemplowska, "Making Justice Sensitive to Responsibility"; Oberdiek, "Perfecting Distributive Justice".

members of society, even if the offenders themselves benefit from mitigation of punishment, doing so will reduce the *total* well-being of worst off members of society. These two objections do not pose a big threat to the empirical claim of the improvement of the worst off argument *if* it is true that *generally speaking*, reducing the punishment of a non-socially deprived offender will not improve the well-being of any worst off members of society. But that does not seem to me to be obviously true. For example, consider the policy of reducing the punishment by a small amount for offenders who have donated a lot of money to charity. This mitigating policy, which of course mainly benefits non-socially deprived offenders, may induce rich people to donate, and as a result, may improve the well-being of the worst off people in general. Accordingly, since it is not true that for all socially deprived offenders, reducing their punishment will improve the well-being of the worst off members of society, and for some non-socially deprived offenders, reducing their punishment will lead to that consequence, I think we cannot accept the empirical claim of the improvement of the worst off argument without further empirical evidence.

The question I now want to ask is: does the whole life moral desert argument, through focusing on a different consequence (namely whole life moral desert rather than distributive justice) avoid the difficulties faced by the improvement of the worst off argument?

I think the answer is no as the empirical claim of the whole life desert argument is vulnerable to basically the same difficulties as the empirical claim of the improvement of the worst off argument. First, while socially deprived offenders are likely to be people who deserve less suffering based on the idea of whole life moral desert, punishing them less does not always reduce their long term suffering; second, punishing them less may sometimes increase the suffering on other people who also deserve less suffering based on whole life moral desert;[414] third, while punishing non-socially deprived offenders less will compromise whole life moral desert *if we focus on the direct effect of sentence on the offender alone*, sometimes a mitigating policy that mainly benefits non-socially deprived offenders will promote whole life moral desert through indirect means (say, by inducing rich people to donate to the suffering people to make the suffering people happier and hence promotes whole life moral desert).

---

414 Ezorsky, "The Ethics of Punishment", xxvii.

Moreover, I believe the moral claim of the whole life moral desert argument faces *extra* difficulties that are *not* faced by the moral claim of the improvement of the worst off argument. To see this point, let us take a look at the *structure* of the moral claim of the improvement of the worst off argument. The moral claim of the improvement of the worst off argument can be broken down into three sub-claims. First, a claim about *value*: improvement of the worst off is valuable due to distributive justice. Second, a claim about the legitimate concern of the *state:* that distributive justice is a value that the state should care about. Third, a claim about the legitimate concern of *sentencing institutions*: that distributive justice is a value that sentencing institutions should care about. I think the first two sub-claims are quite plausible and the only real difficulty lies in defending the third sub-claim (a difficulty that I tried to overcome in chapter 8 above).

Similarly, the moral claim of the whole life moral desert argument can be broken down into three sub-claims. For the whole life moral desert argument, its sub-claim about the legitimate concern of sentencing institutions is similar to the corresponding sub-claim of the improvement of the worst off argument and thus they face similar difficulties. The sub-claim about value and the sub-claim about the legitimate concern of the state of the whole life moral desert argument, however, are less plausible than the corresponding sub-claims of the improvement of the worst off argument. Many people doubt whether whole life moral desert is really a valuable state of affairs but I think fewer people will doubt whether distributive justice is valuable.[415] Many people doubt whether promoting whole life moral desert is a legitimate concern of the state but I think the claim that a state should care about distributive justice, even if not universally accepted, is much less controversial.[416]

3. Klein's payment in advance argument

Klein offered a non-consequentialist non-exculpatory non-communicative argument for claiming that there is a mitigating factor that is more applicable to socially deprived offenders.[417] According to Klein, while an offender should pay for his crime by suffering for it, if his crime was caused by a painful process, then the painful process can be regarded as "payment in advance" for the crime. Klein then observes that the mental states of socially deprived offenders, which motivated their crimes, are more likely to be acquired through a more painful process. To illustrate, let us start by thinking about a socially deprived offender who attacked another person. According

---

[415] Scanlon, "Punishment and the Rule of Law", 220-221.
[416] Ross, "Punishment."
[417] One interpretive point. While Klein talks about "blameworthiness", I think it is clear that she meant to discuss the deservingness of hard treatment rather than the deservingness of reactive attitudes *per se*. See Klein, *Determinism, Blameworthiness, and Deprivation,* 175. Hence, I believe it is appropriate to regard her argument as a non-exculpatory argument, given how I define blame (see chapter 1 above).

to Klein, it might well be the case that "the causal process which has culminated in his becoming someone who lacks feelings for others is one which has involved suffering for him."[418] Now think about a rich offender who attacked another person and who grew up in a luxurious environment. The violent motive of this rich offender may be attributable to the fact that he was spoiled in the past, and being spoiled, while not something good, is not something painful. Accordingly, since socially deprived offenders are more likely to have paid (at least partly) for their crimes in advance, there is a mitigating factor that is more applicable to them.[419] Let me lay out this argument, which can be called the *payment in advance argument*, in a schematic manner before assessing it:

The Payment in Advance Argument

Moral claim: If the *causal process* that led to an offender's crime was more painful, then there is a reason to punish him less, even if he is equally blameworthy for his crime and even if our court's right to communicate blame to him is not compromised. (Let me call this the *payment in advance principle*.)

[418] Klein, *Determinism, Blameworthiness, and Deprivation*, 82.
[419] Klein, *Determinism, Blameworthiness, and Deprivation*, 82.

Empirical claim: On average, the causal processes that led to the crimes by socially deprived offenders are more painful.

Now let me assess the payment in advance argument. I wish to grant, for the sake of discussion, the empirical claim of the argument and I will focus on the payment in advance principle. Before I assess the payment in advance principle, let me stress that the payment in advance principle is *not the same* as a principle that holds that we should count the offender's suffering *in his whole life* and impose less punishment on him if he has suffered more in his life in total. That principle, which is related to the whole life moral desert idea, does *not* have a *causal requirement*: it counts *all* suffering *regardless* of whether the suffering *caused* the offender's crime. In contrast, it is a *distinctive* feature of the payment in advance principle that it counts *only some*, but *not all*, of an offender's suffering in his life: it counts *only* the suffering *that caused the crime*, but not others. Klein has emphasised this causal requirement explicitly.[420]

What then are the reasons to count some, but not all, suffering of the offender in deciding the amount of punishment? Klein has provided three justifications for it but I

[420] Klein, *Determinism, Blameworthiness, and Deprivation*, 81-82.

think all are unsound. The first justification Klein offers for the payment in advance principle is simply that it accounts for people's intuitive judgments concerning punishment of wrongdoers.[421] This justification is not entirely helpful. I think many people simply do not share the intuition that the suffering of the offender matters to how much punishment he should receive. Moreover, even for people who believe that suffering matters to punishment (including me), I doubt if their intuition is one that incorporates the causal requirement.

The second justification offered by Klein for the payment in advance principle, which is more indirect, is that a principle like the payment in advance principle is plausible concerning *rewards for good deeds*, so analogously, we should accept the payment in advance principle in punishment for wrongdoings. Klein invites us to reflect on two equally well-performed students.[422] The first student is naturally talented. He is lazy, but since he is so bright he did well in spite of his laziness. The second student was dumb. His good performance is a result of hard work, and hard work is a painful process. According to Klein, the second student deserves, if not more marks, some extra reward and he deserves more reward than the first student, no matter how much suffering or happiness the two students otherwise had in their lives. That suggests,

[421] Klein, *Determinism, Blameworthiness, and Deprivation*, 81-82.
[422] Klein, *Determinism, Blameworthiness, and Deprivation*, 168-169.

Klein argues, the amount of reward we impose for a good deed should depend on whether the causal process leading to the good deed was a painful one. Analogously, the amount of punishment we impose for a wrongdoing should depend on whether the causal process leading to it was a painful one.

In reply, some people may claim that they do not feel that the dumb student should be rewarded more. Some people may, alternatively, argue that principles governing rewards and principles governing punishments are unrelated so whatever principle of reward we derive from the student example, it is not applicable to punishment. I want to, instead, offer a third line of critique. I do not think, even if we assume that the dumb student should be rewarded more, the student example above supports the principle that if a person's good deed was caused by a more painful process then he should receive more reward.

Let us grant, for the sake of argument, the judgment that the dumb student should be rewarded more. But in order to explain that judgement, we do not need to rely on the principle that if a person's good deed was caused by a more painful process then he should receive more reward. We can explain it by a different principle, which can be called the *effort principle*: a person should receive more reward if he puts more effort in achieving the good deed. The effort principle is, for two reasons, different from the principle that if a person's good deed was caused by a more painful process then he should receive more reward. First, in order for some suffering to count as effort, the agent must *intend* the suffering to contribute to the achievement of the good deed. If my suffering in the past caused a good deed but I had no idea that my suffering would have that result, then it is implausible to say that I made extra effort for the good deed. Second, the effort principle only looks at the amount of effort made by an agent for the good deed without looking at whether the effort *actually contributed to the good deed causally*.

The dumb student did not simply suffer more in the process of achieving the good result; he made extra effort to achieve the good result. So the effort principle, in addition to the principle that if a person's good deed was caused by a more painful process then he should receive more reward, can justify the conclusion that the dumb student should be rewarded more. The effort principle, moreover, is the more plausible principle. This can be illustrated by thinking about the following third student. The third student performed well in the exam, and he worked hard as he thought he was dumb and needed to work hard in order to perform well. However, hard work would not actually help him much for this exam (contrary to what he thought). He was, unknown to himself, really smart, and his good performance was entirely caused by his intelligence and not by his hard work. My intuition is that this third student deserves equal reward as the dumb student. This intuition is consistent with the effort principle but it is inconsistent with the principle that if a person's good deed was caused by a more painful process then he should receive more reward.

The third justification for the payment in advance principle offered by Klein appeals to the idea of distributive justice.[423] But while distributive justice will care about the suffering of the offender, it is unclear on what basis distributive justice will insist on a causal requirement, i.e. it is unclear that distributive justice will count only his suffering that *caused* his crime but not his other suffering.

Based on the above, I think neither our intuition concerning punishment for wrongdoings, nor our intuition concerning rewards for good deeds, nor the idea of distributive justice, justifies the payment in advance principle.

4. Conclusion

[423] Klein, *Determinism, Blameworthiness, and Deprivation*, 169.

In this chapter, I have examined two more non-exculpatory non-communicative arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders.

The first argument examined in this chapter is the whole life moral desert argument, i.e. the argument that claims that there is a reason to reduce the punishment of an offender for his crime if that will promote whole life moral desert, and for a greater proportion of socially deprived offenders than non-socially deprived offenders, reducing their punishment will promote whole life moral desert. It is a consequentialist argument like the improvement of the worst off argument, which I have examined in chapter 8. My assessment of the argument is that the empirical claim of the whole life moral desert argument faces similar challenges as those faced by the empirical claim of the improvement of the worst off argument. Moreover, I have argued that the moral claim of the whole life moral desert argument is less plausible than the moral claim of the improvement of the worst off argument.

The second argument examined in this chapter is the payment in advance argument, i.e. the argument that claims that there is a reason to punish an offender less for his crime if the causal process that led to his crime was a more painful one, and on average the causal processes that led to the crimes of socially deprived offenders are more painful. I have surveyed three justifications for the moral claim of the payment in advance argument and I have argued that they are all unsound.

Chapter 11: Concluding chapter

In the previous nine chapters, I have explored twelve arguments for claiming that there is a mitigating factor that is more applicable to socially deprived offenders. For the sake of conciseness, I will not rehearse all the main points raised in my thesis. Instead, I will point to five areas, which are not covered in this thesis and where further contribution can be made.

First, I have not considered arguments that are based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence. I appreciate that I have excluded a significant category of arguments. The reason for excluding those arguments is, as I mentioned in chapter 1, simply limitation of space. I hope to consider those arguments on another occasion.

Second, I made no claim to comprehensiveness, *even* in relation to arguments that are not based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence. While I have considered quite a number of arguments, I am open to the possibility that some plausible arguments are not covered. Further research can be done to cover other arguments. But I hope this thesis has at least covered most of the prominent arguments in the literature (that are not based on the idea that socially deprived offenders are more likely to act under some mental disorder or disturbance or other forms of irrationality at the time of offence) and can therefore, provide a reference for future work in this area.

Third, more work needs to be done on the comparative weight between different values. I have briefly discussed some all things considered claims, which involve comparison of different values, in this thesis. For example, I have argued that the values that would be compromised by the mitigation of punishment, like proportionality, do not enjoy a general priority over the value of improving the well-being of the worst off so in many cases we should mitigate an offender's punishment to promote distributive justice, all things considered. But clearly, that is all stated at a very general level. In order to decide, in a particular case, whether we should mitigate the punishment of an offender for his crime and by how much, all things considered, we need a more precise way to compare the values. It would be asking for too much, I believe, to look for a universal algorithm; but it would be overly pessimistic to believe that further theoretical guidance is impossible.

Fourth, more empirical research is needed. Partly it is because it is impossible to settle all things considered issues without empirical research. Moreover, further empirical research is also important because for two arguments, namely, the motivating strength argument and the improvement of the worst off argument, I have argued that we do not have sufficient empirical evidence at present to decide whether the empirical claims of those arguments are true.

Fifth, closely related to the last two issues, assuming that we agree, in principle, we ought to reduce the punishment of some offenders for their crimes we still need to consider *how* exactly the mitigating plea ought to be implemented in practice. First, there is an issue about the degree of mitigation. Second, we should also consider whether it should be done by just giving the judges a sentencing discretion, or allowing the legislature to provide rigid sentencing guidelines. I cannot settle these issues here. I only wish to remark that there is no reason to believe that we can find a one-size-fit-all solution: there is no reason to believe that the best proposal for one place will share exactly the same form as the best proposal in another place.

Bibliography (only includes materials that are referred to)

I. Books and articles

Abrams, D. (2012) "Estimating the Deterrent Effect of Incarceration using Sentencing Enhancements", *American Economic Journal: Applied Economics* 4(4): 32

Adams, R. (1985) "Involuntary Sins", *Philosophical Review* 94(1): 3

Alexander, L. (2000) "Insufficient Concern: A Unified Conception of Criminal Culpability", *California Law Review* 88(3): 931

Alexander, L. and Ferzan, K. (2009) *Crime and Culpability* (Cambridge: Cambridge University Press)

Anand, S. (2000) "The Sentencing of Aboriginal Offenders, Continued Confusion and Persisting Problems: A Comment on the Decision in R. v. Gladue", *Canadian Journal of Criminology* 42: 412

Anderson, E. (1999) "What is the Point of Equality?", *Ethics* 109(2): 287

Aquinas, T. (2006) *Summa Theologica: vol. 38: Injustice* (Trans: Marcus Lefébure. Cambridge: Cambridge University Press)

Argyle, M. (1994) *The Psychology of Social Class* (London: Routledge)

Aristotle. *Nicomachean Ethics* (Trans: W. D. Ross. url:

http://classics.mit.edu/Aristotle/nicomachaen.7.vii.html)

Arneson, R. (1990) “Against Complex Equality”, *Public Affairs Quarterly* 4(2): 99

Arneson, R. (2000) “Luck Egalitarianism and Prioritarianism”, *Ethics* 110(2): 339

Arneson, R. (2000) “Two Cheers for Capabilities”, in Brighouse, H. and Robeyns, I. (eds), *Measuring Justice: Primary Goods and Capabilities* (Cambridge: Cambridge University Press)

Arneson, R. (2000) “Welfare should be the Currency of Justice”, *Canadian Journal of Philosophy* 30(4): 497

Arneson, R. (2005) “Distributive Justice and Basic Capability Equality: ‘Good Enough’ Is Not Good Enough”, in Kaufman, A. (ed), *Capabilities Equality: Basic Issues and Problems* (London: Routledge)

Arneson, R. (2011) “Luck Egalitarianism – A Primer”, in Knight, C. and Stemplowska, Z. (eds), *Responsibility and Distributive Justice* (Oxford: Oxford University Press)

Arpaly, N. (2003) *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford: Oxford University Press)

Arpaly, N. and Schroeder, T. (1999) “Praise, Blame, and the Whole Self”, *Philosophical Studies* 93(2): 161

Ashworth, A. (1981) “The Elasticity of Mens Rea”, in Tapper, C. (ed), *Crime, Proof, and Punishment*: *Essays in Memory of Sir Rupert Cross* (London: Butterworths)

Ashworth, A. (1984) "Justifying the Grounds of Mitigation", *Criminal Justice Ethics* 13: 5

Ashworth, A. (2010) *Sentencing and Criminal Justice: 5th ed* (Oxford: Oxford University Press)

Ashworth, A. and Horder, J. (2013) *Principles of Criminal Law: 7th ed* (Oxford: Oxford University Press)

Baron, S., Kennedy, L., and Forde, D. (2001) "Male Street Youths' Conflict: The Role of Background, Subcultural, and Situational factors", *Justice Quarterly* 18(4): 759

Barry, B. (1984) "Intimations of Justice", *Columbia Law Review* 84(3): 806

Barry, B. (1991) "Chance, Choice, and Justice", in his *Essays on Political Theory: Liberty and Justice* (Oxford: Clarendon Press)

Bauer, M. (2013) "How to Avoid the Temptations of Immediate Gratification", *Scientific American*. url: http://www.scientificamerican.com/article/how-to-avoid-the-temptations-of-immediate-gratification/

Bayern, S. (2009) "The Significance of Private Burdens and Lost Benefits for a Fair-play Theory of Punishment", *New Criminal Law Review* 12: 1

Bazelon, D. (1988) *Questioning Authority* (New York: Knopf)

Bell, B., Jaitman, L., and Machin, S. (2014) "Crime Deterrence: Evidence From the

London 2011 Riots", *The Economic Journal* 124(576): 480

Bennett, C. (2004) "The Limits of Mercy", *Ratio* 17(1): 1

Benson, P. (2010) "Contract", in Patterson, D. (ed) *A Companion to Philosophy of Law and Legal Theory: 2nd ed* (Oxford: Wiley)

Bentham, J. (1907) *An Introduction to the Principles of Morals and Legislation* (Oxford: Clarendon Press)

Berman, M. (2008) "Punishment and Justification", *Ethics* 118(2): 258

Berman, M. (2011) "Two Types of Retributivism", in Green, S. and Duff, A. (eds), *Philosophical Foundations of Criminal Law* (Oxford: Oxford University Press)

Boonin, D. (2008) *The Problem of Punishment* (Cambridge: Cambridge University Press)

Brandt, R. (1992) *Morality, Utilitarianism, and Rights* (Cambridge: Cambridge University Press)

Bratman, M. (2000) "Fischer and Ravizza on Moral Responsibility and History", *Philosophy and Phenomenological Research* 61(2): 453

Brink, D. (2013) "Responsibility, Incompetence, and Psychopathy", *Lindley Lecture* 53: 1

Brodeur, J. and Roberts, J. (2002) "Taking Justice Seriously", *Canadian Criminal Law Review* 7: 77

Bronstein, J., Buccafusco, C., and Masur, J. (2009) "Happiness and Punishment", *The University of Chicago Law Review* 76: 1037

Brown, D. (2002) “Third Party Interests in Criminal Law”, *Texas Law Review* 80: 1383

Brown, D. (2004) “Cost-benefit Analysis in the Criminal Law”, *California Law Review* 92(2): 323

Burke, J. (2003) “Reinventing Contract”, *Murdoch University Electronic Journal of Law* 10(2)

Buss, S. (1987) “Justified Wrongdoing”, *Nous* 31(3): 337

Butler, P. (2000) “Class-based Remedies for the Poor”, in Heffernan, W. and Kleinig, J. (eds), *From Social Justice to Criminal Justice* (Oxford: Oxford University Press)

Cahill, M. (2011) “Punishment Pluralism”, in White, M. (ed), *Retributivism: Essays on Theory and Policy* (Oxford: Oxford University Press)

Campbell, C. A. (1951) “Is Free Will a Pseudo-problem?”, *Mind* 60(240): 441

Card, C. (1972) “On Mercy”, *Philosophical Review* 81(2): 182

Carter, M. (2002) “Of Fairness and Faulkner”, *Saskatchewan Law Review* 65*: 63*

Casal, P. (2007) “Sufficiency is not Good Enough”, *Ethics* 117(2): 296

Choo, A. (2008) *Abuse of Process and Judicial Stays of Criminal Proceedings: 2nd ed* (Oxford: Oxford University Press)

Ciurria, M. (2014) “Moral Responsibility: Justifying Strawson and the Excuse of Peculiarly Unfortunate Formative Circumstances”, *Ethical Theory and Moral Practice* 17(3): 545

Clear, T. (2008) “The Effects of High Imprisonment Rates on Communities”, *Crime* *and Justice* 37(1): 97

Cohen, G. A. (1989) "The Currency of Egalitarian Justice", *Ethics* 99(4): 906

Cohen, G. A. (1995) *Self-Ownership, Freedom, and Equality* (Cambridge: Cambridge University Press)

Cohen, G. A. (2006) "Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists?", *Royal Institute of Philosophy Supplement* 58: 113

Cohen, G. A. (2011) "Fairness and Legitimacy in Justice", in his *On the Currency of Egalitarian Justice and Other Essays in Political Theory* (Ed: Otsuka, M. New Jersey: Princeton University Press)

Cohen, G. A. (2011) "Freedom and Money", in his *On the Currency of Egalitarian Justice and Other Essays in Political Theory* (Ed: Otsuka, M. New Jersey: Princeton University Press)

Comfort, M. (2008) *Doing Time Together: Love and Family in the Shadow of the Prison* (Chicago: Chicago University Press)

Cryer, R. and Simester, A. (2006) "Iraq and the Use of Force: Do the Side-effects Justify the Means", *Theoretical Inquiries in Law* 7: 9

Cudjoe, J. and Barringer, T. (2002) "More than Mere Ripples: The Interwoven Complexity of Female Incarceration and the African-American Family", *University of Maryland Law Journal of Race, Religion, Gender and Class* 2(2): 265

Dagan, H. (2013) "Autonomy, Pluralism, and Contract Law Theory", *Law and*

Contemporary Problems 76: 19

Davidson, D. (1980) *Essays on Actions and Events* (Oxford: Oxford University Press)

Day, J. P. (1993) “Temptation”, *American Philosophical Quarterly* 30: 175

Delgado, R. (1985) “‘Rotten Social Background’: Should the Law Recognize a Defence of Severe Environmental Deprivation?’, *Law and Inequality* 3: 9

Dolinko, D. (1991) “Some Thoughts about Retributivism”, *Ethics* 101(3): 537

Dolinko, D. (2006) “Some Naïve Thoughts about Justice and Mercy”, *Ohio State Journal of Criminal Law* 4: 349

Doob, A. and Webster, C. (2003) “Sentence Severity and Crime: Accepting the Null Hypothesis”, *Crime and Justice* 30: 143

Duff, A. (1986) *Trials and Punishments* (Cambridge: Cambridge University Press)

Duff, A. (1990) *Intention, Agency, and Criminal Liability* (Oxford: Oxford University Press)

Duff, A. (1998) “Principle and Contradiction in the Criminal Law”, in Duff, A. (ed), *Philosophy and the Criminal Law* (Cambridge: Cambridge University Press,)

Duff, A. (2000) “Whose Law is it Anyway? Inclusion, Exclusion and the Criminal Law”, *Scottish Affairs* 30: 2

Duff, A. (2001) *Punishment, Communication, and Community* (Oxford: Oxford University Press)

Duff, A. (2003) “Virtue, Vice, and Criminal Liability: Do We Want an Aristotelian

Criminal Law?", *Buffalo Criminal Law Review* 6: 147

Duff, A. (2007) "The Intrusion of Mercy", *Ohio State Journal of Criminal Law* 4: 361

Duff, A. (2007) *Answering for Crime* (Oxford: Hart Publishing)

Duff, A. (2010) "A Criminal Law for Citizens", *Theoretical Criminology* 14: 293

Duff, A. (2010) "Authority and Responsibility in International Criminal Law", in Besson, S. and Tasioulas, J. (eds), *The Philosophy of International Law* (Oxford: Oxford University Press)

Duff, A. (2010) "Blame, Moral Standing, and the Legitimacy of the Criminal Trial", *Ratio* 23: 123

Dworkin, R. (2002) *Sovereign Virtue* (Cambridge: Harvard University Press)

Ebels-Duggan, K. (2013) "Dealing with the Past: Responsibility and Personal history", *Philosophical Studies* 164(1): 141

Editors of the Yale Law Journal (1977) "On Prisoners and Parenting: Preserving the Tie that Binds", *Yale Law Journal* 84: 1408

Ezorsky, G. (1977) "The Ethics of Punishment", in Ezorsky, G. (ed), *Perspectives on Punishment* (Albany: State University of New York Press)

Fagan, J., West, V., and Holland, J. (2003) "Reciprocal Effects of Crime and Incarceration in New York City Neighborhoods", *Fordham Urban Law Journal* 30: 1551

Falk, P. (1995) "Novel Theories of Criminal Defense Based upon the Toxicity of the Social Environment: Urban Psychosis, Television Intoxication, and Black

Rage", *North Carolina Law Review* 74: 732

Feinberg, J. (1974) "Noncomparative justice", *Philosophical Review* 83(3): 297

Feinberg, J. (1974) "The Expressive Function of Punishment", in his *Doing and Deserving* (New Jersey: Princeton University Press)

Feinberg, J. (1986) *Harm to Self* (Oxford: Oxford University Press)

Fields, L. (1994) "Moral Beliefs and Blameworthiness", *Philosophy* 69(270): 397

Finnis, J. (2011) *Natural Law and Natural Rights: 2nd ed* (Oxford: Oxford University Press)

Fischer, J. and Ravizza, M. (1994) "Introduction", in Fischer, J. and Ravizza, M. (eds), *Perspectives on Moral Responsibility* (Ithaca: Cornell University Press)

Fischer, J. and Ravizza, M. (1998) *Responsibility and Control*: *A Theory of Moral Responsibility* (Cambridge: Cambridge University Press)

Frank, N. and Groves, B. (1986) "Punishment, Privilege, and Structured Choice", in Groves, B. and Newman, G. (eds), *Punishment and Privilege* (New York: Harrow and Heston)

Frankfurt, H. (1988) "Alternative Possibilities and Moral Responsibility", in his *The Importance of What We Care About* (Cambridge: Cambridge University Press)

Frankfurt, H. (1988) "Coercion and Moral Responsibility", in his *The Importance of What We Care About* (Cambridge: Cambridge University Press)

Frankfurt, H. (1988) "Equality as a Moral Ideal", in his *The Importance of What We Care About* (Cambridge: Cambridge University Press)

Frankfurt, H. (2002) "Reply to Fischer", in Buss, S. and Overton, L. (eds), *Contours* *of Agency: Essays on Themes from Harry Frankfurt* (Cambridge: MIT Press)

Franklin, E. (2013) "A Theory of the Normative Force of Pleas", *Philosophical Studies* 163(2): 479

Gardner, J. (2007) "In Defence of Defences", in his *Offences and Defences* (Oxford: Oxford University Press)

Gardner, J. (2007) "The Gist of Excuses", in his *Offences and Defences* (Oxford: Oxford University Press)

Garvey, S. (2009) "Dealing with Wayward Desire", *Criminal Law and Philosophy* 3: 1

Garvey, S. (forthcoming) "Injustice, Authority, and the Criminal Law", in Sarat, A. (ed), *The Punitive Imagination* (Alabama: The University of Alabama Press)

Goldman, A. "Toward a New Theory of Punishment", *Law and Philosophy* 1(1): 57

Green, L. (2010) "Legal Obligation and Authority", in Edward Zalta (ed), *Stanford Encyclopedia of Philosophy*, url: http://plato.stanford.edu/entries/legal-obligation/

Green, S. (2009) "Hard Times, Hard Time: Retributive Justice for Unjustly Deprived Offenders", *University of Chicago Legal Forum*: 43

Green, T. H. (1999) *Lectures on the Principles of Political Obligation* (Kitchner: Batoche Books)

Griffin, J. (1988) *Well-being* (Oxford: Oxford University Press)

Grogger, J. (1995) “The Effect of Arrests on the Employment and Earnings of Young Men”, *The Quarterly Journal of Economics* 110(1): 51

Haque, A. (2013) “Revolution and the Criminal Law”, *Criminal Law and Philosophy* 7(2): 231

Harris, M. (1987) “A Brief for De-escalating Criminal Sanctions”, in Gottfredson, S. and McConville, S. (eds), *American's Correctional Crisis: Prison Populations and Public Policy* (Westport: Greenwood)

Hart, H. L. A. (2007) “Legal Responsibility and Excuses”, in his *Punishment and Responsibility: 2nd ed* (Ed: Gardner, J. Oxford: Oxford University Press)

Hart, H. L. A. (2007) “Prolegomenon to the Principles of Punishment”, in his *Punishment and Responsibility: 2nd ed* (Ed: Gardner, J. Oxford: Oxford University Press)

Hestevold, S. (1985) “Justice to Mercy”, *Philosophy and Phenomenological Research* 46(2): 281

Hieronymi, P. (2004) “The Force and Fairness of Blame”, *Philosophical Perspectives* 18: 115

Hild, M. and Voorhoeve, A. (2004) “Equality of Opportunity and Opportunity Dominance”, *Economics and Philosophy* 20: 117

Hoffman, P. (2005) “Aquinas on Threats and Temptations”, *Pacific Philosophical Quarterly* 86: 225

Holroyd, J. (2010) "Punishment and Justice", *Social Theory and Practice* 36(1): 78

Holton, R. (1999) "Intention and Weakness of Will", *Journal of Philosophy* 96(5): 241

Holton, R. (2003) "How is Strength of Will Possible?", in Tappolet, C. and Stroud, S. (eds), *Weakness of Will and Practical Irrationality* (Oxford: Oxford University Press)

Holton, R. and Shute, S. (2007) "Self-control in the Modern Provocation Defence", *Oxford Journal of Legal Studies* 27(1): 49

Holtug, N. and Lippert-Rasmussen, K. (2006) "An Introduction to Contemporary Egalitarianism", in Holtug, N. and Lippert-Rasmussen, K. (eds), *Egalitarianism: New Essays on the Nature and Value of Equality* (Oxford: Clarendon Press)

Horder, J. (1994) "Rethinking Non-fatal Offences against the Person", *Oxford Journal of Legal Studies* 14(3): 335

Horder, J. (2004) *Excusing Crime* (Oxford: Oxford University Press)

Horne, T. (1981) "Review of *A Discourse on Property*", *Political Theory* 9(3): 451

Hsieh, N. (2007) "Incommensurable Values", in Edward Zalta (ed), *Stanford Encyclopedia of Philosophy*, url: http://plato.stanford.edu/entries/legal-obligation/

Hudson, B. (1993) *Penal Policy and Social Justice* (London: Palgrave Macmillan)

Hudson, B. (1994) "Beyond Proportionate Punishment: Difficult Cases and the 1991 Criminal Justice Act", *Crime, Law, and Social Change* 22: 59

Hudson, B. (1999) "Punishment, Poverty, and Responsibility", *Social and Legal Studies* 8: 583

Hudson, B. (2000) "Punishing the Poor", in Heffernan, W. and Kleinig, J. (eds), *From Social Justice to Criminal Justice* (Oxford: Oxford University Press)

Hurka, T. (2003) "Desert: Holistic and individualistic", in Serena Olsaretti (ed), *Desert and Justice* (Oxford: Oxford University Press)

Husak, D. (1998) "Partial Defences", *Canadian Journal of Law and Jurisprudence* 11: 167

Husak, D. (2011) "Negligence, Belief, Blame and Criminal Liability: The Special Case of Forgetting", *Criminal Law and Philosophy* 5: 199

Husak, D. (2012) "'Broad' Culpability and the Retributivist Dream", *Ohio State Journal of Criminal Law* 9: 449

Husak, D. (2013) "Retributivism in Extremis", *Law and Philosophy* 32: 3

Johnson, C. (1991) "Entitled to Clemency: Mercy in the Criminal Law", *Law and Philosophy* 10(1): 109

Jones, P. (2010) "Group Rights", in Edward Zalta (ed), *Stanford Encyclopedia of Philosophy*, url: http://plato.stanford.edu/entries/legal-obligation/

Jung, H. (2011) "Increase in the Length of Incarceration and the Subsequent Labor Market Outcomes: Evidence from Men Released from Illinois State Prisons", *Journal of Policy Analysis and Management* 30(3): 499

Kant, I. (1931), *Lectures on Ethics* (Trans: Louis Infield. New York: Century Company)

Karmen, A. (2000) "Poverty, Crime, and Criminal Justice", in Heffernan, W. and

Kleinig, J. (eds), *From Social Justice to Criminal Justice* (Oxford: Oxford University Press)

Kaye, A. (2005) “Resurrecting the Causal Theory of the Excuses”, *Nebraska Law Review* 83(4): 1116

Kaye, A. (2007) “The Secret Politics of the Compatibilist Criminal Law”, *Kansas Law Review* 55: 365

Kelly, E. (2013) “Culpability and Fault in Criminal Justice” (A paper presented at the Chinese University of Hong Kong on 19 June 2013)

Keren-Paz, T. (2007) *Tort, Egalitarianism, and Distributive Justice* (Surrey: Ashgate)

Khoury, A. (2013) “Synchronic and Diachronic Responsibility”, *Philosophical Studies* 165(3): 735

King, M. (2008) “The Structure of Responsibility” (PhD thesis submitted to the University of Maryland)

Klein, M. (1990) *Determinism, Blameworthiness and Deprivation* (Oxford: Oxford University Press)

Kleinig, J. (1991) “Punishment and Moral Seriousness”, *Israel Law Review* 25: 401

Kling, J. (2006) “Incarceration Length, Employment and Earnings”, *American Economic Review* 96(3): 863

Kordana, A. and Tabachnick, D. (2005) “Rawls and Contract Law”, *George Washington Law Review* 73(3): 598

Kymlicka, W. (2002) *Contemporary Political Philosophy: 2nd ed* (Oxford: Oxford

University Press)

Landerso, R. (2012) *Does Incarceration Length Affect Labor Market Outcomes for Violent Offenders?* (Copenhagen: University Press of Southern Denmark)

Leverick, F. and Chalmers, J. (2008) "Fair Labelling in Criminal Law", *Modern Law Review* 71(2): 217

Levy, N. (2003) "Cultural Membership and Moral Responsibility", *The Monist* 86(2): 145

Levy, N. (2011) *Hard Luck* (Oxford: Oxford University Press)

Levy-Pounds, N. (2007) "From the Frying Pan into the Fire: How Poor Women of Color and Children are Affected by Sentencing Guidelines and Mandatory Minimums", *Santa Clara Law Review* 47(2): 285

Lewis, C. (2011) "Temptation, Stratification, and Proportional Punishment" (A paper presented at the Ockham Society, University of Oxford, on 1st February 2011)

Lippke, R. (2003) "Diminished Opportunities, Diminished Capacities", *Social Theory and Practice* 29(3): 459

Lippke, R. (2007) *Rethinking Imprisonment* (Oxford: Oxford University Press)

Lippke, R. (2011) "Social Deprivation as Tempting Fate", *Criminal Law and Philosophy* 5: 277

Lipsey, M. and Cullen, F. (2007) "The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews", *Annual Review of Law and Social Science*

3: 297

Litton, P. (2005) “The ‘Abuse Excuse’ in Capital Sentencing Trials: Is it Relevant to Responsibility, Punishment, or Neither?”, *American Criminal Law Review* 42: 1027

Loeffler, C. (2013) “Does Imprisonment Alter the Life Course? Evidence on Crime and Employment from a Natural Experiment” *Criminology* 51(1): 137

Lopez-Calva, L., Rigolini, J., and Torche, F. (2011) “Is There Such Thing as Middle Class Values?: Class Differences, Values and Political Orientations in Latin America”, The World Bank Policy Research Paper Series 5874

Lyons, D. (1986) “Coercion as Temptation”, *Journal of Social Philosophy* 17: 35

Maitland, F. W. (1880) “The Relation of Punishment to Temptations”, *Mind* 5: 259

Mann, S. and Al-Khadha, M. (2004) “Freedom of the Will and Criminal Culpability”, *University of New South Wales Law Review* 8*:* 98

Markowitz, F. (2003) “Socioeconomic Disadvantage and Violence”, Aggression and Violent Behavior 8(2): 145

Markowitz, F. and Felson, R. (1998) “Social-demographic Attitudes and Violence”, Criminology 36(1): 117

Maskin, E. (1995) “Correlation and Constraint”, *Boston Review* 20(2)

Mason, A. (2000) “Equality, Individual Responsibility, and Gender Socialization”, *Proceedings of the Aristotelian Society* 100: 227

Matravers, M. (2006) “Who’s Still Standing: A Comment on Antony Duff’s Precondition of Criminal Liability”, *Journal of Moral Philosophy* 3(3): 320

McCoy, C. (1997) “Sentencing (and) the Underclass”, *Law and Society Review* 31(3): 589

McKenna, M. (2004) “Responsibility and Globally Manipulated Agents”, *Philosophical Topics* 32(1): 169

McKenna, M. (2008) “A Hard-line Reply to Pereboom’s Four-case Manipulation Argument”, *Philosophy and Phenomenological Research* 77(1): 142

McKenna, M. (2012) *Conversation and Responsibility* (Oxford: Oxford University Press)

Melissaris, M. (2012) “Property Crimes as Crimes of Injustice”, *Criminal Law and Philosophy* 6: 149

Metz, T. (2007) “How to Reconcile Liberal Politics with Retributive Punishment”, *Oxford Journal of Legal Studies* 27(4): 683

Milo, R. (1983) “Wickedness”, *American Philosophical Quarterly* 20(1): 69

Mitchell, B. (2001) “Multiple Wrongdoing and Offence Structure” *Modern Law Review* 64(3): 393

Montmarquet, J. (1992) “Epistemic Virtue and Doxastic Responsibility”, *American Philosophical Quarterly* 29(4): 331

Montmarquet, J. (2007) “‘Pure’ versus ‘Practical’ Epistemic Justification”, *Metaphilosophy* 38(1): 71

Moore, M. (2010) "Causation and the Excuses", in his *Placing Blame* (Oxford: Oxford University Press)

Moore, M. (2010) "The Moral Worth of Retribution", in his *Placing Blame* (Oxford: Oxford University Press)

Morris, H. (1968) "Persons and Punishment", *The Monist* 52(4): 475

Morris, N. (1974) "The Future of Imprisonment", *Michigan Law Review* 72(6): 1161

Morse, S. (2000) "Deprivation and Desert", in Heffernan, W. and Kleinig, J. (eds), *From Social Justice to Criminal Justice* (Oxford: Oxford University Press)

Morse, S. (2011) "Severe Environmental Deprivation: A Tragedy, Not a Defence", *Alabama Civil Rights and Civil Liberties Law Review* 2: 147

Murphy, J. (1994) "Marxism and Retribution", in Duff, A. and Garland, D. (eds), *A Reader on Punishment* (Oxford: Oxford University Press)

Murphy, L. (2003) *Moral Demands in Nonideal Theory* (Oxford: Oxford University Press)

Murray, J. and Farrington, D. (2008) "The Effects of Parental Imprisonment on Children", *Crime and Justice* 37(1): 133

Nagin, D., Cullen, F., and Jonson, C. (2009) "Imprisonment and Reoffending", *Crime and Justice* 38(1): 115

Nelkin, D. (forthcoming) "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness", *Nous*

Norrie, A. (2002) "From Criminal Law to Legal Theory: The Mysterious Case of the Reasonable Glue Sniffer", *Modern Law Review* 65: 538

Nozick, R. (1974) *Anarchy, State, and Utopia* (New York: Basic Books)

Nussbaum, M. (1993) "Equity and Mercy", *Philosophy and Public Affairs* 22(2): 83

Oberdiek, J. (2014) "Perfecting Distributive Justice", url: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2501444

Odudu, O. (2003) "Retributive Justice in an Unjust Society", *Ratio Juris* 16: 416

Ormerod, D. (2011) *Smith and Hogan's Criminal Law: 13th ed* (Oxford: Oxford University Press)

Parent, W. D. (1976) "The Whole Life View of Criminal Desert", *Ethics* 86(4): 350

Parfit, D. (1991) "Equality or Priority", *Lindley Lecture*: 1

Parfit, D. (2012) "Another Defence of the Priority View", *Utilitas* 24(3): 399

Perry, S. (1997) "Risk, Harm, and Responsibility", in Owen, D. (ed) *The Philosophical Foundations of Tort Law* (Oxford: Clarendon Press)

Piqueri, A. and Blumstein , A. (2007) "Does Incapacitation Reduce Crime", *Journal of Quantitative Criminology* 23(4): 267

Rawls, J. (1999) *A Theory of Justice: revised ed* (Oxford: Oxford University Press)

Raz, J. (1985) *The Morality of Freedom* (Oxford: Oxford University Press)

Reimen, J. (2007) "The Moral Ambivalence of Crime in an Unjust Society", *Criminal Justice Ethics* 26: 3

Richards, N. (1987) “Acting under Duress”, The *Philosophical Quarterly* 37: 21

Risse, M. (2002) “What Equality of Opportunity Could Not be”, *Ethics* 112(4): 720

Ristroph, A. (2006) “Desert, Democracy, and Sentencing Reform”, *Journal of Criminal Law and Criminology* 96(4): 1293

Roach, K. and Rudins, J. (2002) “Broken Promises: A Response to Stenning and Roberts’ ‘Empty Promises”’, *Saskatchewan Law Review* 65: 3

Roberts, D. (2000) “The Ethics of Punishing Indigent Parents”, in Heffernan, W. and Kleinig, J. (eds), *From Social Justice to Criminal Justice* (Oxford: Oxford University Press)

Roberts, J. (2002) “The Sentencing of Aboriginal Offenders in Canada: A Rejoinder”, *Saskatchewan Law Review* 65(1): 75

Robichaud, P. (forthcoming) “On Culpable Ignorance and Akrasia”, *Ethics* 125(1): 137

Robinson, P. (1982) “Criminal Law Defences: a Systematic Analysis”, *Columbia Law Review* 82: 199

Robinson, P. (2013) “Are we Responsible for Who we Are? The Challenge for Criminal Law Theory in the Defenses of Coercive Indoctrination and ‘Rotten Social Background’”, *Alabama Civil Rights and Civil Liberties Review* 2: 53

Roemer, J. (1995) “Equality and Opportunity”, *Boston Review* 20(2)

Roemer, J. (1995) “John Roemer Replies”, *Boston Review* 20(2)

Roemer, J. (2000) *Equality of Opportunity* (Cambridge: Harvard University Press)

Roemer, J. (2003) "Defending Equality of Opportunity", *The Monist* 86(2): 261

Rosen, G. (2003) "Culpability and Ignorance", *Proceedings of the Aristotelian Society* 103: 61

Rosen, G. (2004) "Skepticism about Moral Responsibility", *Philosophical Perspectives* 18: 295

Ross, W. D. (1977) "Punishment", in Ezorsky, G. (ed), *Perspectives on Punishment* (Albany: State University of New York Press)

Ryberg, J. (2004) *The Ethics of Proportionate Punishment: A Critical Investigation* (London: Kluwer)

Sadurski, W. (1985) "Distributive Justice and the Theory of Punishment", *Oxford Journal of Legal Studies* 5(1): 47

Scanlon, T. (1986) "The Significance of Choice", *The Tanner Lectures on Human Values* 7: 149

Scanlon, T. (2003) "Punishment and the Rule of Law", in his *The Difficulty of Tolerance* (Cambridge: Cambridge University Press)

Scanlon, T. (2003) *What We Owe to Each Other* (Cambridge: Harvard University Press)

Scanlon, T. (2008) *Moral Dimensions* (Cambridge: Harvard University Press)

Scheffler, S. (2003) “What is Egalitarianism?”, *Philosophy and Public Affairs* 31(1): 5

Schoeman, F. (1987) “Statistical Norms and Moral Attributions”, in Schoeman, F. (ed), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (Cambridge: Cambridge University Press)

Scott, E. (2006) “Adolescent and the Regulation of Youth Crime”, *Temple Law Review* 79: 337.

Scott, E. and Steinberg, L. (2003) “Blaming Youth”, *Texas Law Review* 81(3): 838

Sen, A. and Williams, B. (1982) “Introduction”, in Sen, A. and Williams, B. (eds), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press)

Shafer-Landau, R. (2003) *Moral Realism*: *a Defence* (Oxford: Oxford University Press)

Shelby, T. (2007) “Justice, Deviance, and the Dark Ghetto”, *Philosophy and Public Affairs* 35(2): 126

Sher, G. (1987) *Desert* (Oxford: Oxford University Press)

Shoemaker, D. (2007) “Moral Address, Moral Responsibility, and the Boundaries of the Moral Community”, *Ethics* 118(1): 70

Shoemaker, D. (2012) “Blame and Punishment”, in Coates, J. and Tognazzini, N. (eds), *Blame: Its Nature and Norms* (Oxford: Oxford University Press)

Simester, A. (1996) “Why Distinguish Intention from Foresight”, in Simester, A. and Smith, A. (eds), *Crime and Culpability* (Oxford: Oxford University Press)

Simester, A. (2009) “Intoxication is Never a Defence”, *Criminal Law Review*: 3

Simester, A. and von Hirsch, A. (2011) *Crimes, Harms, and Wrongs* (Oxford: Hart

Publishing)

Simmons, A. J. (1979) *Moral Principles and Political Obligation* (New Jersey: Princeton University Press)

Simmons, A. J. (2001) "The Obligations of Citizens", in his *Justification and Legitimacy* (Cambridge: Cambridge University Press)

Simmons, A. J. (2010) "Ideal and Non-ideal Theory", *Philosophy and Public Affairs* 38: 5

Smart, A. (1968) "Mercy", *Philosophy* 43(166): 345

Smith, A. (2005) "Responsibility for Attitudes: Activity and Passivity in Mental Life", *Ethics* 115(2): 236

Smith, A. (2007) "On Being Responsible and Holding Responsible", *The Journal of Ethics* 11: 465

Smith, A. (2008) "Control, Responsibility, and Moral Assessment", *Philosophical Studies* 138: 367

Smith, A. (2012) "Moral Blame and Moral Protest", in Coates, J. and Tognazzini, N. (eds), *Blame: Its Nature and Norms* (Oxford: Oxford University Press)

Smith, H. (1983) "Culpable Ignorance", *Philosophical Review* 92(4): 543

Song, L. and Lieb, R. (1993) *Recidivism: The Effect of Incarceration and Length of Time Served* (Olympia: Washington State Institute for Public Policy)

Stemplowska, Z. (2009) "Making Justice Sensitive to Responsibility", *Political Studies* 57: 237

Strawson, P. (1993) “Freedom and Resentment”, in Fischer, J. and Ravizza, M. (eds), *Perspectives on Moral Responsibility* (Ithaca, Cornell University Press)

Sunstein, C. (2007) “If People Would Be Outraged by Their Rulings, Should Judges Care?”, *Stanford Law Review* 60: 155

Tadros, V. (2005) *Criminal Responsibility* (Oxford: Oxford University Press)

Tadros, V. (2009) “Poverty and Criminal Responsibility”, *Journal of Value Inquiry* 43: 391

Tadros, V. (2011) *The Ends of Harm* (Oxford: Oxford University Press)

Talbert, M. (2009) “Implanted Desires, Self-formation, and Blame”, *Journal of Ethics and Social Philosophy* 3(2)

Talbert, M. (2012) “Moral Competence, Moral Blame, and Protest”, *Journal of Ethics* 16: 89

Talbert, M. (2013) “Unwitting Wrongdoers and the Role of Disagreement in Blame”, in Shoemaker, D. (ed) *Oxford Studies in Agency and Responsibility vol 1* (Oxford: Oxford University Press)

Tasioulas, J. (2003) “Mercy”, *Proceedings of the Aristotelian Society* 103: 101

Temkin, L. (2002) “Equality, Priority, and the Leveling Down Objection”, in Clayton, M. and Williams, A. (eds), *The Ideals of Equality* (London: Palgrave MacMillan)

Temkin, L. (2011) “Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck", in Knight, C. and Stemplowska, Z. (eds), *Responsibility and Distributive Justice* (Oxford: Oxford University Press)

Tonry, M. (1994) "Proportionality, Parsimony, and Interchangeability of Punishments", in Duff, A. and Garland, D. (eds), *A Reader on Punishment* (Oxford: Oxford University Press)

Tonry, M. (2013) "Sentencing in America: 1975-2025", *Crime and Justice* 42: 141

UK Home Office (1990) *Crime, Justice, and Protecting the Public* Cm 965 (London: Her Majesty's Stationery Office)

US Bureau of Justice Statistics. (2003) *Prevalence of Imprisonment in the US population, 1974-2001* (Washington, DC: Government Publication Office)

Vargas, M. (2005) "The Trouble with Tracing", *Midwest Studies in Philosophy* 29(1): 269

Vincent, N. (2008) "Book Review of *Torts, Egalitarianism and Distributive Justice*", *Australasian Journal of Legal Philosophy* 33: 199

Vogel, L. (1993) "Understanding and Blaming", *Philosophy and Phenomenological Research* 53(1): 129

Von Hirsch, A. (1976) *Doing Justice: The Choice of Punishments* (New York: Hill and Wang)

Von Hirsch, A. (1992) "Proportionality in the Philosophy of Punishment", *Crime and Justice* 16: 55

Von Hirsch, A. (1994) "Sentencing Reform", in Duff, A., Marshall, S., Dobash, R., and Dobash, R. (eds), *Penal Theory and Practice* (Manchester: Manchester University Press)

Von Hirsch, A. (1996) *Censure and Sanctions* (Oxford: Oxford University Press)

Von Hirsch, A. and Ashworth, A. (2005) *Proportionate Sentencing: Exploring the Principles* (Oxford: Oxford University Press)

Voorhoeve, A. (2007) *Equal Opportunity, Equality, and Responsibility* (PhD thesis submitted to the University of London)

Voorhoeve, A. (2008) "Scanlon on Substantive Responsibility", *Journal of Political Philosophy* 16(2): 184

Vuoso, G. (1987) "Background, Responsibility, and Excuse", *Yale Law Journal* 96*:* 1661

Wakefield, S. and Wilderman, C. (2013) *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality* (Oxford: Oxford University Press)

Waldfogel, J. (1994) "The Effect of Criminal Conviction on Income and the Trust Reposes in the Workmen", *The Journal of Human Resources* 29(1): 62

Waldron, J. (2013) "'To Bestow Stability upon Possession': Hume's Alternative to Locke", in Penner, J. and Smith, H. (eds), *Philosophical Foundations of Property Law* (Oxford: Oxford University Press)

Wallace, R. J. (1998) *Responsibility and the Moral Sentiments* (Oxford: Oxford University Press)

Wallace, R. J. (2010) "Hypocrisy, Moral Address, and the Equal Standing of Persons", *Philosophy and Public Affairs* 38(4): 307

Walzer, M. (1983) *Spheres of Justice: A Defence of Pluralism and Equality* (New York: Basic Books)

Watson, G. (2004) "Disordered Appetites: Addiction, Compulsion, and Dependence", in his *Agency and Answerability: Selected Essays* (Oxford: Oxford University Press)

Watson, G. (2004) "Free Agency", in his *Agency and Answerability: Selected Essays* (Oxford: Oxford University Press)

Watson, G. (2004) "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme", in his *Agency and Answerability: Selected Essays* (Oxford: Oxford University Press)

Wellman, C. (2001) "Toward a Liberal Theory of Political Obligation", *Ethics* 111(4): 735

Westen, P. (2006) "An Attitudinal Theory of Excuse", *Law and Philosophy* 25(3): 289

Western, B. and Pettit, B. (2010) *Collateral Costs: Incarceration's Effect on Economic Mobility* (Washington: Pew Charitable Trusts)

Western, B., Kling, J., and Weiman, D. (2001) "The Labor Market Consequences of

Incarceration", *Crime and Delinquency* 47(3): 410

White, M. (2011) "Pro Tanto Retributivism: Judgment and the Balance of Principles in Criminal Justice", in White, M. (ed), *Retributivism: Essays on Theory and Policy* (Oxford: Oxford University Press).

Wilderman, C. (2009) "Parental Imprisonment, the Prison Boom, and the Concentration of Childhood Disadvantage", *Demography* 46(2): 265

Williams, A. (2006) "Liberty, Liability, and Contractualism", in Holtug, N. and Lippert-Rasmussen, K. (eds), *Egalitarianism: New Essays on the Nature and Value of Equality* (Oxford: Oxford University Press)

Williams, B. (1995) "Internal Reasons and the Obscurity of Blame", in his *Making Sense of Humanity: and Other Philosophical Papers, 1982-1993* (Cambridge: Cambridge University Press)

Williams, R. (2012) "Voluntary Intoxication – a Lost Cause?", *Law Quarterly Review* 126: 264

Wolf, S. (1987) "Sanity and the Metaphysics of Responsibility", in Schoeman, F. (ed), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (Cambridge: Cambridge University Press)

Wolf, S. (1990) *Freedom within Reason* (Oxford: Oxford University Press)

Wolff, J. (1991) *Robert Nozick: Property, Justice, and the Minimal State* (Stanford: Stanford University Press)

Wright, R. G. (1996) *Does the Law Morally Bind the Poor? Or, What Good's the Constitution When You Can't Afford a Loaf of Bread?* (New York: New

York University Press)

Zimmerman, M. (1997) “Moral Responsibility and Ignorance”, *Ethics* 107(3): 410

II. English statutes

Criminal Damage Act 1971 s 1

Criminal Justice Act 2003 ss 114, 142

The Rehabilitation of Offenders Act 1974

III. English cases

Dao [2012] EWCA Crim 1717

Kingston [1994] 3 WLR 519

Majewski [1976] UKHL 2

R v C [2013] EWCA Crim 223

Re A (Conjoined Twins) [2001] 2 WLR 480

Southwark LBC v Williams [1971] Ch 734

IV. Statutes from other jurisdictions

(Canada) *Criminal Code* s 718.2(e)

(Hong Kong) *Crimes Ordinance* s 137

V. Cases from other jurisdictions

(Australia) *Bugmy* [2013] HCA 37

(Australia) *O'Connor* (1980) 29 ALR 44

(Canada) *Gladue* [1999] 1 S.C.R. 688

(Canada) *Ipeelee* [2012] 1 S.C.R. 433

(US) *Penry*, 492 U.S 319 (1989)