

Blame-Based Expressivism: Explaining the Permissibility of State Punishment Using Scanlonian Blame

Joanna Mary Firth

Keble College, University of Oxford
D.Phil, Department of Politics and International Relations

Contents Page

<i>Acknowledgments</i>	<i>i</i>
<i>Abstract</i>	<i>ii</i>
<i>Introduction</i>	<i>1</i>
 Chapter 1	
<i>What's Wrong with other Justificatory Theories?</i>	<i>20</i>
 Chapter 2	
<i>Previous Expressive and Rights Forfeiture Theories</i>	<i>57</i>
 Chapter 3	
<i>Blame-Based Expressivism</i>	<i>92</i>
 Chapter 4	
<i>The State's Standing to Blame</i>	<i>114</i>
 Chapter 5	
<i>Which Rights May the State Suspend From Criminal Wrongdoers?</i>	<i>143</i>
 Chapter 6	
<i>Problems with Thinking of Punishment as a Response to Blameworthiness</i>	<i>194</i>
 <i>Concluding Remarks</i>	<i>227</i>
<i>Bibliography</i>	<i>254</i>

Acknowledgments

I would like to thank the following people and institutions, none of whom can be blamed for any failings in this work. To David Miller, my supervisor. This thesis has benefited enormously from his guidance and insights, for which I am very grateful. To the Arts and Humanities Research Council, the Society of Applied Philosophy, and the University of Oxford's Vice Chancellor's Fund, for the research grants and scholarships that supported this thesis. I had the privilege of participating in and presenting at Oxford's Political Theory Graduate Seminar, the Nuffield Political Theory Workshop, and the Manchester Centre for Political Theory. Thanks to all members for lively discussions. To David Birks, Ian Carrol, Iason Gabriel, Sarah Hannan, Jeff Howard, Hugh Lazenby, Rob Munro, Rebecca Reilly-Cooper, and Juri Viehoff for numerous debates and fond memories of my time as a graduate student. To my current boss, Paul Sutton, for granting me leave to complete the final stages of this thesis, and to my colleagues at Greig City Academy Science Department for covering my classes during my absence. To the children I've taught over the last three years for allowing me to return to my thesis with renewed enthusiasm, better perspective, and greater depth to add to my examples. To all my own teachers along the way, with special thanks to Jonathan Quong for pushing me to apply for the D. Phil, a great deal of encouragement, and many, many discussions and comments. To him I owe a huge debt of gratitude. To my all friends for moral support and to Katie Bilboa, Alan Lockey, Cameron Corbett, Louise Gilmer, and Laura Graham extra thanks are due for much appreciated proof reading and comments. Cameron Corbett deserves special appreciation for his love and support in this final stretch of writing. Finally, I would like to thank my parents, to whom this thesis is dedicated, for their amazing generosity and kindness, without which my thesis would not have been possible.

Abstract

This thesis argues that it is all-things-considered permissible for the state to punish citizens who commit criminal wrongs. It does this by combining elements of expressive and rights forfeiture theories of punishment. It argues that when one citizen criminally wrongs another, this damages her relationship with that citizen and the citizenry-at-large. Citizens have special citizenship rights against each other, grounded on the fact that they stand in a special relationship with each other. Thus when a criminal wrongdoer damages this relationship, the state (acting on behalf of the citizenry-at-large) is licensed to suspend their special citizenship rights from them. There is positive value in it suspending these rights because it expresses moral criticism of the wrongdoer. I argue the right to roam freely around a state's territory and the right to vote are citizenship rights which can be suspended. I argue that prison and vote removal therefore are permissible forms of punishment. According to my theory, however, there are strict limits on the rights that can be suspended from wrongdoers: there is no justification for removing rights that are not grounded on the special citizenship relationship, such as basic rights to lead a minimally decent life. This has implications for the types of prison that can be permitted. My thesis contributes to the literature in three main ways: First, it develops and defends two initially appealing types of theory – expressivist and rights forfeiture - which have nonetheless been dismissed by many philosophers. Second, it is unusual in playing explicit attention to the important question of which forms punishment can permissibly take. Third, it discusses what conditions in prison should be like. This is an area of great moral importance, but it is chronically under-researched by philosophers.

Introduction

Many people support the idea that the state may punish citizens who are guilty of criminal wrongdoing. State punishment comes in different forms. Examples include Home Detention Curfew, community service, fines, and prison. Yet, whatever form of punishment a state uses, it involves inflicting burdensome sanctions on criminal wrongdoers and sending them the message that it disapproves of their conduct. Thus whatever form it takes, state punishment conflicts with our moral rules about how a state should ordinarily treat its citizens. When it puts wrongdoers on Home Detention Curfew, it monitors their movements through electronic tagging and confines them to their homes for certain prescribed hours each day. When the state puts wrongdoers on community service, it forces them to take part in often onerous activities that they probably do not wish to take part in. When the state fines wrongdoers, it takes money off them beyond what it takes off other citizens of similar wealth in taxes. When the state imprisons wrongdoing citizens, it undermines their autonomy by exercising huge control over their daily routines and the longer-term options they can pursue, it isolates them from the social and economic spheres of our society, and it separates them from their loved ones.

Given the way that punishment conflicts with our moral rules about how a state should ordinarily treat its citizens, we need to be able to explain why it is morally permissible. We need to be able to explain this in a general sense: why does the fact that a citizen has committed a crime make it permissible for the state to criticise her and burden her in a way that would ordinarily be impermissible? We also need to be able to explain the

permissibility of punishment in a more particular sense: we need to be able to explain what types – what concrete modes - of punishment are permissible. This thesis defends the permissibility of state punishment in the general sense (henceforth simply ‘punishment’). It also provides an explanation of the type of punishments permitted. It does not consider all possible types of punishment, but one of the types it does consider is prison and it argues that these can be permissible. Of course it does not argue that all prisons as they currently are are permissible. Many of our current prisons treat inmates in ways that are not permissible – they do not, for example, allow inmates adequate opportunities to maintain family relationships or spend time out of their cell. Rather my thesis defends a more liberal form of prison. My thesis also argues that removing the right to vote is a permissible form of punishment.

I believe that removing the vote is an intuitive form of punishment and I believe it is desirable that a theory of punishment be able to explain the permissibility of prison. In my mind removing the vote seems a fitting response to criminal wrongdoers. These people have violated our society’s laws and it is intuitive to think that removing something that symbolises being part of that society is appropriate. I believe that it is desirable to be able to explain the permissibility of prison for two reasons. First, in Britain today prison is the first thing that springs to mind at the mention of state punishment, even though we use and have used other forms. Not only is it foremost in our thoughts, but it is also extremely widely used, with an estimated 80 thousand people behind bars in England and Wales (Ministry of Justice, 2016). Therefore if we are unable to explain its permissibility this would require a radical shift in the way we think about and carry out state punishment. Of course, if it turns out that prison truly is impermissible, then we have to make this radical

shift.¹ But if we were to abandon prison as a form of punishment, we would need to come up with an alternative. This would to some degree mean taking a step into the unknown in an area of life which is very important to us as a society. As such, until such an alternative has been presented and carefully thought through, it seems advantageous to be able to explain the permissibility of prison.

Second, there is something attractive about the idea of prison as a form of punishment, at least when it comes to relatively serious crimes. Although their original purpose was to prevent the escape of those awaiting trial, their use as a form of punishment proper mirrors the Athenian practice of exile in the message it conveys. It conveys the message that we do not want those who commit criminal wrongs to move freely amongst us. In the case of serious wrongs, I think this is appealing. When a citizen commits a serious criminal wrong, it seems right for other citizens to feel distanced from her and one way this can be manifested is by temporarily and/or partially separating her from them through putting her in prison. Many theorists mention prison as the main form of punishment in society in passing, and thus probably assume that their theories can justify the practice. But not many theorists explicitly consider this assumption and, as my discussion below will indicate, they might be incorrect to make it. Thus in tackling this question directly, my thesis provides a contribution to the literature.

¹ This may in the end be the right thing to do; there is important recent philosophical work on abolitionism (Boonin, 2008) and some prison charities offer pragmatic arguments in favour of limited abolition. For example, Francis Crook of the Howards League for Penal Reform argues for the abolition of women's prisons largely because of the effects of separating women from their children (Crook, 2006).

I. My Justification of Punishment and its Main Contribution to the Literature

Many different types of theory have been put forward with the intention of explaining why punishment is justified, but my thesis combines elements of the expressive and rights forfeiture traditions. Expressivists seek to justify punishment on the basis that it seems right that the state may or should express moral criticism of wrongdoers. This is their foundational claim. They then make the further claim that this criticism may take the form of punishment. The expressivist starting point that the state should or may criticise wrongdoers is deeply attractive. If, as a society, we have decided that a certain action is wrong and should be criminalised, then it seems correct to say that our state may, all else being equal, criticise those who perform that action. It also seems good for the state to express criticism: it is important both because, in cases where there is a clearly identifiable victim, it shows him that we care about the wrong he has suffered and because it affirms the society's values and preserves a sense of solidarity amongst citizens.

However, previous expressive theories of punishment suffer from a major problem. This is that the further claim they make does not hold: there are many ways to express criticism of wrongdoing. Therefore the conclusion that the state may express criticism of wrongdoing through *punishment* (either in the general sense or in the sense of prison) does not follow from the fact that it may or should express criticism of wrongdoing. Why not, express it in some other, less damaging way? Why not, as T.M. Scanlon famously suggested, 'say it with flowers?' (Scanlon, 1988: 214). Expressive theorists need to be able to specify the ways in which the state may express criticism of wrongdoing and explain why punishment

is permissibly included within this set of actions. Expressive theorists have made many attempts to do this, none of which, I will argue, are successful.

Rights forfeiture views punishment draw attention to a crucial insight: in order for it to be permissible to harm a person, usually it must be the case that she has no right against this harm. Punishment is harmful. Therefore, in order for it to be permissible to punish a person, it usually² must be the case that the crime she has committed caused her to forfeit her right against that particular punishment. This point is undoubtedly true. Thus any theory of punishment must take account of this crucial insight: it must be able to explain why wrongdoing citizens lack a right against what the state is doing to her in the name of punishment (but see Lipkke, 2007). However, rights forfeiture views as they have traditionally been interpreted (henceforth ‘traditional right forfeiture views’) suffer from a problem. They have been interpreted as implying a sort of *strict reciprocity*: they have been interpreted as implying that the right that the wrongdoer forfeits is the same or an equivalent right as the one he violated when committing his crime. So, for example, if one citizen rapes another, rights forfeiture views have been interpreted as implying that he loses right against being raped (or an ‘equivalent right’). But we clearly this is a deeply implausible implication as would not want our state to punish wrongdoers this way.

In defence of rights forfeiture views, it is important to note that, though some theorists have explicitly argued for strict reciprocity (e.g. Goldman, 1979), rights forfeiture does not necessarily imply it. There is, to be sure, an appealing symmetry to the idea that the right

² I say ‘usually’ so as not to rule out the possibility that there can be (rare) circumstances in which it is permissible to punish a person even though this would infringe her rights. This could happen if, for example, punishing her would somehow if prevent some major catastrophe or we knew that she was highly dangerous and certain to commit multiple terrible offences.

violated by the wrongdoer should be the right that is forfeited, but it is by no means a logical entailment. From the idea that a person forfeits some of her rights by committing a criminal wrong, nothing particular logically follows about *which* rights these are (c.f. Wellman, 2012: 386-387). Nonetheless, to be plausibly complete, a rights forfeiture theorist must say something about which particular rights are lost. So, just like the expressivists, rights forfeiture theorists must find a way of starting to specify the ways in which the state may treat wrongdoers. Not many rights forfeiture theorists have done this (but see McDermott, 2001), so there is a gap in the literature that needs to be addressed. Another issue rights forfeiture theorists need to contend with is that showing wrongdoers have forfeited their rights against punishment is not enough to show that punishment is all-things-considered permissible. Punishment is expensive and it has a negative effect on wrongdoers and their families, so we also need to show that there is positive value to punishment that can explain why these costs are worth paying.

Many philosophers have written off expressive theories as justifications of punishment on the basis that they do not seem to be able to explain why the state may or should adopt punishment instead of other forms of moral criticism. Many philosophers have also written off all rights forfeiture views on the basis of the problem that only applies to the traditional interpretation of that view: that the implications of strict reciprocity are deeply implausible. But, I think that both expressive and rights forfeiture views contain some very important ideas: it is good to express criticism of wrongdoers and we must be able to explain why wrongdoers lack rights against punishment. Thus, the chief purpose of my thesis is to argue that we can combine elements of these views to give an attractive justification of state punishment, whilst avoiding some of their typical pitfalls. This is the

main way I intend my thesis to contribute to the literature. Here follows a summary of its main argument.

I call my theory *blame-based expressivism*. This I because I believe that the type of moral criticism we should use T.M. Scanlon's conception of blame as the basis of our explanation as to why state punishment is permissible. When we blame a person, this involves judging that she is *blameworthy*. Judging that a person is blameworthy means judging that her moral deliberations about why to pursue an action were *morally faulty*. But this is not all there is to blame – it also involves a shift in our attitude towards the person we have judged to be blameworthy. As Coates and Tognazzini say, '[w]hen we blame someone, we're not simply *noting* the fact that she falls below some standard; rather, blame seems to also be about our own attitudes toward' that agent (Coates and Tognazzini, 2013: 10). There are debates around how best to understand the content of this shift in attitude. But Scanlon argues that, as well as making the judgment that a person is blameworthy for a wrongful action, blaming them is 'to take that action to indicate something about the person that impairs one's relationship with him or her, and to understand that relationship in a way that reflects this impairment' (Scanlon, 2008: 123). So the shift in attitude combines the judgement that a person is blameworthy with an attitude that reflects the idea that she has impaired her relationship with you. This could include a general feeling of distance from the person, a decreased interest in seeing them and/or a decrease in inclination to be friendly or supportive to them (Scanlon, 2013: 11, 2008: 143).

Scanlon intends his account of blame to provide a fairly comprehensive account of what blame is. However, as a comprehensive account of blame Scanlon's view faces problems – most notably that it seems difficult for it to account for blame between strangers (Sher, 2013: 50-55) and that it seems to lack some of the more passionate emotions or attitudes, such as anger or resentment, that sometimes seem to form part of blame (Smith, 2013: 38; Mason, 2011; Wallace, 2011: 350). Thus I should state from the outset that I am not using Scanlon's view as a comprehensive account of blame. What I want to take from Scanlon's account is this: within the context of special relationships, if one party wrongs the other this results in relationship impairment. If the wronged party makes the judgement that the wrongdoer is blameworthy and alters her attitude and behaviour towards him in such a way that reflects this, this is a form of blame. I only claim, however, that this is just one type of blaming and it is to this type of blame that I am referring to when I say 'blame' or 'Scanlonian blame'.

Adopting the Scanlonian conception of blame as a basis for state punishment is useful because it gives us a framework through which we can specify the ways in which the state may express criticism of wrongdoing citizens. If a relationship is impaired when one person wrongs another, then, absent overriding considerations such as that some disproportionately bad consequence will ensue, I believe this means that special rights that were grounded in this relationship may be withdrawn or suspended. This is because the source of those special rights (the special relationship) is damaged. However, it does not follow that rights which were not grounded in this special relationship can be suspended. This is because the source of these rights has not been damaged.

Citizens stand in a special relationship with one another – that of co-citizenship. This relationship grounds special citizenship rights. *Citizenship rights* are special rights that we hold only in virtue of being a citizen of a state and not simply in virtue of us being persons. They arise out of our special relationship with other citizens. I argue that when one citizen wrongs another this damages their special co-citizenship relationship with the rest of the citizens in the state. This means that the state may suspend some of the special citizenship rights the wrongdoer possessed. I argue that citizenship rights include the right to vote and, more controversially, the right to move freely around the state's territory. If we accept these claims, then we can see why my theory gives us a way of starting to specify why the state may express criticism of wrongdoers through (practices that we would recognise as) punishment. This is because removing the right to vote is a well-recognised form of punishment and, if I am right that the right to move freely around the territory is a citizenship right and can thus be suspended from wrongdoers, we have the beginnings of an argument as to why the state may imprison wrongdoers (since imprisonment necessarily involves some level of confinement).

However, blame-based expressivism sets firm limits on the rights that can be withdrawn from wrongdoers. This is because, as well as having rights grounded in our special relationships, we have some rights just in virtue of the fact that we are persons. One category of these rights is *basic rights* which are rights to have what we need so that we can live a minimally decent life. Since these rights do not arise out of special relationships, damaging our relationships does not damage the source of these rights and thus blame-based expressivism does not permit the state to withhold them as a form of punishment.³

³ I do not believe that it is impossible to forfeit basic rights. I believe for example, that it is possible to forfeit your right against being killed or seriously harmed if you culpably attack an innocent

The fact that blame-based expressivism sets these limits on the way that the state may punish wrongdoers is an attractive feature of the theory. The idea that even the worst criminals retain basic rights is an intuitive one, and a plausible theory of punishment must be able to explain it (or at least be consistent with it). The fact that blame-based expressivism cannot permit the removal of wrongdoer's basic rights clearly has implications for what the conditions inside prison should be like. I believe that we have basic rights to some level of social inclusion, freedom to see loved ones, autonomy, and having our basic physical needs met are basic rights. This explains why I argue for a fairly liberal form of prison. Philosophers are mostly agreed that current prisons are unjust. Thus, considering how many people are currently in prison, it seems a matter of urgency that we consider what a permissible prison would look like. However, this is something of an untouched research area inside the philosophy of punishment (but see Lippke, 2007), so the fact that my theory considers this issue in depth represents an important contribution to the literature.

We can now start to see how blame-based expressivism can overcome the problems of previous expressive theories and traditional rights forfeiture views. Blame-based expressivism breaks free of the implausible strict reciprocity understanding of the rights forfeiture views. It does not imply that rapists should be punished with the removal of the right against rape or equivalent. It claims that rapists undermine their citizenship relationship with their co-citizens and that their citizenship rights can be suspended. It thus narrows down the set of rights wrongdoers forfeit when they commit a crime and thus

person and harming or killing you is the only way in which she can defend herself against death. But *blame-based expressivism* cannot be used to explain the forfeiture of such rights, thus a blame-based expressivist system would not permit the withdrawal of these rights from wrongdoers as a form of punishment.

fills the gap in the rights forfeiture literature. Blame-based expressivism can also overcome the other issue faced by rights forfeiture views – that as well as showing why wrongdoers have forfeited their rights against punishment, they also need to show why it is all-things-considered permissible despite its costs. To do this they need to explain why punishment has positive value. In blame-based expressivism, this positive value comes from the expressive function of punishment. Expressing criticism of wrongdoers is valuable both because it affirms the society's values and preserves a sense of solidarity amongst citizens and because, in cases where there is a clearly identifiable victim, it shows him that we care about the wrong he has suffered. I think that people value these expressive functions of punishment enough to mean that is worth spending our state's resources on it. Blame-based expressivism also overcomes the main problem of previous expressive theories because, by specifying why the right to vote and the right against imprisonment can be lost, it explains why punishment can come within the set of permissible ways in which the state can express criticism of wrongdoers.

If I am right that blame-based expressivism can overcome the problems of previous expressive theories and traditional rights forfeiture views, then this would represent a step forward for both sets of theories. Another way that my thesis contributes to the literature is in its consideration of two other prominent types of theory that seek to justify punishment. These are desert-based retributive and deterrence theories.

Desert-based retributivists seek to justify punishment by appealing to the supposedly widely shared intuition that wrongdoers deserve to suffer. They then use the idea that it is good for wrongdoers to suffer to form part of their justification for state punishment - it is

a means of bringing about the suffering they deserve. *Deterrence theorists* seek to justify punishment on the basis of the crime reduction that it is thought to bring about.

Historically, many deterrence-based justifications of punishment have been consequentialist. Consequentialism is the general moral theory that the permissibility of an action depends *solely* on whether it brings about good consequences. Consequentialists disagree about what counts as a good consequence but, in the context of the debates around the justification of punishment, consequentialists have typically selected the harm reduction brought about by the deterrent effect of punishment as the relevant good consequence (e.g. Smart, 1973). However, one can adopt a deterrence-based justification of punishment without being a consequentialist (Tadros, 2011: 13-14). Another deterrence-based way of seeking to justify punishment is as a form of collective self-defence (Tadros, 2011: 14). There are different versions of this view but they share the following idea: if it is permissible to harm or threaten to harm a wrongdoer to defend yourself from a threat that he or others pose, then this means that it is permissible to harm a wrongdoer through punishment to defend society from the threat of future harms that he or others might pose if we did not have penal institutions. Such views are deterrence-based because it is only permissible to punish offenders if doing so ‘defends’ society by reducing crime (otherwise the harm would be unnecessary). They need not be consequentialist, however, because they need not explain the permissibility of punishment solely in terms of its good consequences – they can also refer to the idea (common in the literature on defensive violence⁴) that by threatening another or by threatening society the wrongdoer or forfeits right against harm.

⁴ See for example (McMahan 2009; Quong, 2009; Thomson, 1991).

However, in my view these theories suffer from serious setbacks. The desert-based retributive idea that wrongdoers deserve to suffer implies that it is non-instrumentally valuable for them to do so. They claim that the intuition that wrongdoer's suffering is a non-instrumentally valuable is widely shared and they rely on this claim. But many people actually find this idea repulsive and the claim that it can be good for a person to suffer is deeply at odds with our normal views on suffering. Thus we need to interrogate whether the intuition that it is good for wrongdoers to suffer is as widely shared as desert-based retributivists claim. I think that they are too quick to assume this. I think they are right that many people have the intuition that we should punish wrongdoers, but this does not show that they also must believe that their suffering is non-instrumentally valuable. Instead, I think our pro-punishment intuitions can often be explained by the expressivist intuition that society should take action to react to wrongdoing. Such action would undoubtedly make wrongdoers suffer, but thinking it is required does not mean we have to endorse the claim that their suffering has non-instrumental value.

The criticisms of consequentialist morality in general and their justifications of punishment in particular are both well known and convincing – the most obvious, in relation to punishment, is that they cannot satisfyingly explain why we ought not to punish the innocent if doing so would bring about the best consequences. On the other hand, collective self-defence views are more sophisticated and have considerably more plausibility. However, I have reservations about these views. The first reservation concerns the implications these views have about the types and magnitude of harm they permit. Collective self-defence views rely on the idea that we may harm people through punishment because we may harm them in self-defence. That is, they rely on the idea that the harm it is permissible to inflict on someone in the defensive realm is transferrable to

the punitive realm. So it would seem that these theories imply that the types and magnitude of harm that are permissible in the defensive context are permissible in the punitive context. But the defensive context frequently permits us to use physical violence and inflict harm of high magnitude. But this is not acceptable in the punitive realm. My second reservation is that, insofar as collective self-defence theorists want to be able to explain the permissibility of *prison*, they place much too much trust in the empirical premise that these institutions do effectively reduce crime. In the empirical literature there is a reasonable amount of consensus that sentence length and type do not have much impact on crime rates. Thus, although they may be able to justify punishment in the general sense it is unlikely that they can justify much use of prison in practice.

II. Overview of the Chapters in the Thesis

The first chapter of the thesis considers desert-based retributive and deterrence theories – and explains my objections to them. The first part of the second chapter is mainly devoted to showing that previous expressive theories fail to explain why the moral criticism of wrongdoers must take the form of hard treatment such as prison. The second part of the chapter explains previous rights forfeiture theories in more detail, and considers and rejects Daniel McDermott's (2001) attempt to specify the types of rights that can be lost. In chapter 3, I give the main positive argument of the thesis. I explain Scanlonian blame, what this means for how we may treat wrongdoers in the contexts of personal relationships, and then I apply this to the context of the relationship between citizens.

In chapter 4 I develop the application of Scanlonian blame to the context of citizenship. Though we can all judge that a wrongdoer is blameworthy, we do not always have the *standing* to actually blame them. This can be the case when the wrongdoing is simply none of our business. Given that most citizens are strangers to each other, we might wonder whether it is all citizens' business when one citizen is wronged. But if we want to ground a theory of punishment *by the state* rather than just by an individual, we need to show that *all* citizens as a collective body, not just the victim, may suspend the wrongdoer's citizenship rights.

Chapter 5 explores the question of which rights are citizenship rights and which rights are basic rights. It defends the claims that the rights to move freely around a territory and to vote are citizenship, rather than basic, rights. It ends by explaining the implications of such rights on what prisons must be like under a system of blame-based expressivism. In chapter 6, I explain why blame-based expressivism is not vulnerable to the typical objections to theories which posit that blameworthiness can form part of the basis of criminal liability. Philosophers have raised a whole host of objections to such theories. These include that such theories are unable to deal with moral luck; that it is inappropriate for the state to make moral assessments of the quality of its citizens' moral reasoning; and that they imply that we may punish citizens for blameworthy things that should not be punished by the state (such as mere thoughts and non-harmful actions, trivial blameworthy acts, and private blameworthy acts).

III. The assumptions behind and further contributions of blame-based expressivism

To really be complete, a theory of punishment needs to be situated within a broader theory of legal and political philosophy. The key issues relevant to punishment are the questions concerning what gives the state *authority* – what gives it authority to act as a collective body on behalf of its people and to claim that citizens have a duty to obey its laws (as well as what then gives it the further authority to punish those who break the laws); what the proper aims of state should be; and the relationship between the rules that determine what makes something a crime (conduct rules) and the set of rules that determine whether, and to what extent, a person should be punished for committing a crime (decision rules). Situating a theory of punishment within a broader theory of legal and political philosophy requires one to work out answers to all these questions. This would be a life's work or more. As such, I need to make some assumptions about some of these questions.

One assumption I make is that the state has the authority to act as a collective body on behalf of its citizens and that it has the authority to declare acts as crimes. I also assume the context of a liberal democracy and that legitimate state action is constrained by the sort of constraints that should limit the actions of a liberal democratic state, for example, that citizens should have privacy in some respects and that they should have a reasonable degree of free thought. This is important in my account because it sets limits on what the state has the standing to punish for. I will not defend it, however, because I assume that liberal democratic values are widely shared and that all theories of punishment must be constrained by these values. In contrast to some other liberal theorists, I assume a conception of the liberal democratic state that sees the state as something more than just an institution to prevent harm or an institution that may *only* appeal to the ideas of freedom

and equality when justifying its actions. I think states should pursue a greater plurality of values including values such as community, solidarity, and recognition.

As regards the relationship between conduct rules and decision rules, I make the assumption that there can be enough separation between the two to allow that a citizen's blameworthiness is relevant to decision rules, but not than conduct rules. I think that crimes are public wrongs, that is, wrongs that are the proper concern of all citizens. In general, citizens will be blameworthy when they perform these acts. But the reason we should make the acts illegal is not that performing them makes a person blameworthy. Nonetheless, if a citizen is blameworthy for performing one of the proscribed actions, her blameworthiness is relevant to why we should punish her.

Although there are questions that we can leave to our broader political or legal theory to answer, a theory of punishment must be able to directly explain (i) why punishment does not violate the rights of wrongdoers; (ii) that there is some positive value to the practice; (iii) that we can reasonably assume the positive value makes the costs permissible; (iv) that punishment actually achieves the positive value that it is meant to; (v) that the punishment permitted by the theory is appropriate in type and severity; and (vi) why the state has the authority to punish (as distinct from the authority to act make laws).

I hope that my above discussion indicates that I think blame-based expressivism can explain points (i) –(vi) But I think it is worthwhile to make a few explicit comments on (v) and (vi). It is tempting to assume that that if the state has the authority to declare acts crimes, then it also has the authority to punish (vii). If this is correct, then it is be

acceptable to leave the question of whether the state has the authority to punish to be answered by our broader political or legal theory. However, this should not be assumed. Some people think that to say 'X is a crime' just is to say that it is an action which is punishable by the state. But this might not be true

As Duff argues, it would make perfect conceptual sense to think of a crime as something that warranted the response of trial and state condemnation but not any further sanction (Duff, 2013). Or we could think of them as condemnable actions that the state may try use force to ensure that citizens do not perform by, for example, using the police force to carry out drug confiscation raids or ensure that restraining orders are followed. But this does not require us to think that a crime is, by definition, an act that is punishable by the state. This is because punishment is not the same as enforcement because it comes after the crime and so does not prevent the particular crime in question (except when punishment is being used as a form of incapacitation).⁵ Therefore, the fact that the state has the authority to declare actions as crimes does not also show that it has the authority to punish. This means that the question of what gives the state the authority to punish might be a question that needs to be answered directly by people who write about this topic. Thus the fact that chapter 4 is mainly concerned with explaining why the state has the standing to punish wrongdoers makes for an interesting contribution to the literature.

In regards to (vi), many theorists comment on whether they think their view has plausible implications for the severity of punishment. But, with the exception of several other

⁵ It has some preventative effect in the sense that it deters citizens would otherwise be tempted to commit crimes from committing them – but the point still remains that we can think of a crime as condemnable actions that the state may try to enforce that citizens do not perform without needing to think that a crime is necessarily something that is punishable by the state.

expressive theories, blame-based expressivism is relatively unusual in paying a lot of attention to the form of punishment it permits. Some of the critical assessments I give of other theories shine a light on the issues that may arise if one does not pay attention to this issue. It is important to consider the forms of punishment your theory permits or else you can be left with unexpected, and perhaps unwanted, implications. For example, if, as I suggested above, collective self-defence views can permit physical harm, this is a defect with those views. Thus a theory of punishment is incomplete, and may be defective, if the form of punishment it permits is not considered.

I hope the arguments of this thesis are successful in revitalising expressivist and rights forfeiture theories of punishment. I think it is beneficial to revitalise these theories both because they contain many attractive elements in their own rights and because I do not think that the main alternative theories (desert-based retributivism and deterrence views) are successful in justifying state punishment. This is the main contribution of the thesis and I hope people will agree with my findings or, if not, I hope that my findings can at least be of use in prompting some further debate. The other main area in which I hope my theory can contribute concerns the form punishment can take. I hope it highlights the need for theorists of punishment to consider this issue more carefully, both because it is important for the overall plausibility of individual theories but also, and more importantly, because we should be giving greater philosophical consideration to the rights of wrongdoers.

CHAPTER 1

What's Wrong With Other Justificatory Theories?

:

In this chapter, I examine the main justifications of punishment from outside the expressive and rights forfeiture traditions and I find them to be wanting. If I am right, and if we want to be able to justify punishment, then this gives us reason to take expressive and rights forfeiture theories of punishment seriously. Justifications of punishment can be divided into two types: retributivist and deterrence-based. In deterrence-based justifications of punishment, the main positive justification of punishment comes from the instrumental value that punishment has *in reducing crime*. In deterrence theories, if punishment does not reduce crime, then there is no positive reason to punish (or at least no positive reason that is thought to do important enough justificatory work to justify very much punishment).

In contrast, retributive justifications posit that punishment has important positive value apart from any value it may have in reducing crime. Desert-based retributivists believe that it is non-instrumentally good for wrongdoers to suffer and therefore that we should intentionally inflict suffering on wrongdoers through punishment. Blame-based expressivism is also a type of retributive view: withdrawing citizenship rights is a permissible response to crime because the relationship between the wrongdoer and the rest of the citizens has become impaired, thus undermining the grounds of the wrongdoers' citizenship rights. Punishment here has positive value apart from any benefits it has in reducing crime because of its expressive function.

I. What's Wrong With Other Types of Retributivism?

Fair-play retributivism

Fair-play theorists believe that punishment is non-instrumentally valuable because it is a form of fairness. Broadly speaking, fair-play views of punishment state that obeying the law is a burden that we must all pay in order to secure the overall benefit that having a system of law brings. Thus, when someone breaks the law she unfairly gains a benefit over the rest of the law-abiding citizens so punishment is justified as a fitting way of cancelling out this benefit (see e.g. Morris, 1968; Sher, 1987; Dagger, 2008).

The fair-play theory sounds plausible when we consider it in the context of crime such as tax evasion or breaking the speed limit because these are cases where the wrongdoer seems to be taking unfair advantage. However, when it comes to more central kinds of crimes it has been subject to decisive objections. These objections are well known so I will only offer a brief outline here. The two most important objections to this view are as follows: (i) for most of us refraining from the most serious types of criminal wrongdoing, such as rape and murder, is not a burden so it is not clear why we should view those who commit such wrongs as having gained a benefit (see e.g. Hampton, 1992: 1660; Dolinko, 1991: 545-546). (ii) Fair-play theories suffer from a 'wrong wrong' problem: Even if rapists and murders did gain a benefit from committing their crimes, it does not really seem like *that* is the appropriate reason to punish them. When we think about why we should respond to

wrongs like rape and murder, any unfair advantage the wrongdoer has gained does not really seem to be important. What matters the action that actually wronged the victim (the rape or the murder): surely *this* is what we should be responding to (see e.g. Duff, 2001, 1986; Hampton, 1992: 1661).

Desert-based Retributivism

Desert based retributivists (DBRs) believe that there is positive non-instrumental value in making wrongdoers suffer. They believe in the *desert-based retributive intuition (DBRI)* that ‘wrongdoers deserve to suffer’.⁶ They believe that it is *good* for wrongdoers to suffer and endorse the further claim that if someone deserves something this means there is a positive moral reason for us to see that he or she gets it. Though they need to endorse the latter claim to explain why humans may bring about punishment, notice that the initial claim, that it is good for wrongdoers to suffer, normally comes prior to this claim as it is thought to be what grounds it. This then would appear to commit DBRs to the idea that it would be good if the suffering of wrongdoers came about without human interference. (Or if it does not we would need an explanation as to why).

The main focus of this section will be explaining the trouble with the idea that it is non-instrumentally valuable for wrongdoers to suffer. However it is worth first pausing to explain why blame-based expressivists are not similarly committed to this idea. In blame

⁶ DBRs often use the language of punishment instead of suffering e.g. ‘wrongdoers deserve to be punished’ rather than ‘wrongdoers deserve to suffer’. However, punishment is conventionally defined as (amongst other things) intentionally making a person suffer. Thus it is a reasonable assumption that DBR’s believe that there is positive non-instrumental value in the *suffering* of wrongdoers.

based-expressivism, wrongdoers *will* suffer (or at least be burdened) when their valuable citizenship rights are withdrawn. But, as Scanlon argues, this is not the same as thinking it is good that the wrongdoers suffer or that it is their suffering that makes withholding their rights appropriate. Withdrawing citizenship rights has become permissible⁷ because the relationship between the wrongdoer and the other citizens has become impaired thus weakening the grounds of the wrongdoers' citizenship rights (Scanlon, 2008: 188). This is similar in structure to the following example: If you reject a romantic suitor, this might cause heartache. But the fact that you reject them can be made appropriate simply by the fact that they are not the right person for you, rather than by any appeal to the idea that the heartache is good (Scanlon, 2008: 189). Thus blame-based expressivists are not committed to the idea that it is non-instrumentally valuable for wrongdoers to suffer. The permissibility of punishment comes from the relationship impairment and the positive reason to punish comes from its expressive function.

Desert-based retributivism has a number of subscribers. Stephen Kershnar and Michael Moore are perhaps the most explicit adherents. Kershnar writes 'deserved punishment is intrinsically valuable' (Kershnar, 2000: 98) and Moore states 'what is distinctively [desert-based] retributivist is the view that the guilty receiving their just deserts is an intrinsic good' (Moore, 1993: 19). Both Kershnar and Moore view retributive desert as providing a weighty, though sometimes overrideable, reason in favour of punishment and believe that this is its main (and in Kershnar's case only) justificatory rationale (Kershnar, 2000: 100).

⁷ At least it has become what I call *prima facie* permissible. Punishment is *prima facie* permissible if it will not violate the rights of wrongdoers. But being *prima facie* permissible does not make it all-things-considered permissible. It may not be all-things-considered permissible if, for example, it is too expensive or if it would violate the rights of the wrongdoers family. I return to this theme in chapter 3.

John Tasioulas and Igor Primoratz each argue for an expressive theory of punishment that is ultimately backed by desert-based retributivism. Primoratz says, for example:

[I]n claiming that punishment is *intrinsically* right and called for [I am committed]...to just those considerations of justice and desert on which the [desert-based] retributivist insists (Primoratz, 1989: 202 his emphasis).

Tasioulas repeatedly claims that punishment is justified because it is deserved. For example, he says:

[O]nly...punishment enables us properly to evince the sincerity of our condemnation [of wrongdoers] *precisely because it is what the offender deserves* (Tasioulas, 2006: 296 my emphasis).

He argues that the fact an offender deserves punishment provides an independent positive reason to punish them (Tasioulas, 2006: 291). Thus he clearly believes there is positive non-instrumental value in inflicting suffering on wrongdoers. Both Primoratz and Tasioulas see it as being important that offenders get what they deserve and thus give retributive desert a fairly central role in their justification of punishment.

It is worth mentioning that some DBR's are more moderate. Douglas Husak, for example, does not believe that retributive desert provides a particularly *weighty* reason in favour of punishment: he believes that desert-based retributivist reasons must almost always be supplemented by deterrent benefits if an act of punishment is to be justified. But, nonetheless, he does (tentatively) subscribe to the view that there is some non-instrumental positive value in making wrongdoers suffer and that this can contribute to the justification of punishment (Husak, 2012: 13-15).

So a reasonable number of philosophers can be classified as being DBRs who appeal to the

DBRI in their work. Some DBR's think it is acceptable to appeal to the DBRI without further justification. Consider Tasioulas:

[R]etributive desert is a basic norm of justice...[it] forms an inextricable part of practices that deeply structure human life, practices which require no justification from a supposed rational standpoint (Tasioulas, 2006: 297, see also Primoratz, 1989).

I disagree. Philosophers need to appeal to intuitions all the time and there is nothing wrong with this *per se*. However, the DBRI is different from other intuitions and we need to be more careful with it. The intuition that we should cause suffering is in stark contrast to what we usually think of as permissible behaviour. One of the least controversial moral intuitions that you can appeal to is the intuition that causing harm or suffering is wrong or bad. Therefore DBRs need to be able to say something to defend their appeals to the DBRI.

One way that DBRs justify their use of the DBRI is to say that it is widely shared (e.g. Husak, 2013: 12, Moore, 1993). For example, Moore says:

[I]magine...Dostoevsky's Russian nobleman in *The Brothers Karamazov*, who turns loose his dogs to tear apart a young boy before the eyes of the boy's mother; imagine further that circumstances are such...that no non-retributive purpose would be served by punishing this offender...[S]hould...the other offender be punished?...*The retributivist 'yes' runs deep for most people* (Moore, 1993: 25 my emphasis).

In the next section I challenge this assumption, however let's assume that it is true for a moment. Saying that there is a broad degree of consensus over an intuition can be an acceptable way of justifying it. But what form does this consensus need to take? It does not need to be universal. We know this because it can be acceptable to appeal to intuitions that not everybody shares. For example not everybody shares the intuition that intended

harm is worse than merely foreseen harm, but philosophers can acceptably appeal to it. But what is not normally the case is that those who lack the intuition actually find it repulsive. For example, if you do not have the intuition that intended harm is worse than foreseen harm, it is likely that you just ‘don’t feel it’ or see the attraction but do not have a strong sense of it, rather than actively thinking it is bad. But there is something different about the lack of consensus over the desert-based retributive intuition. Many people who reject the DBRI do not just ‘not feel it’. They actively find it repulsive, shameful, ‘morally indefensible’ (Scanlon, 2000: 274), or (in some forms) ‘barbaric’ (Tadros, 2011: 62; for similar views see Narveson, 1974: 191-3; Brownlee, 2012:69). Many philosophers think that we should provide more justification of the controversial intuitions than just that they are widely shared. I think the repulsion with which some view the DBRI gives us more reason than normal to require this.

The other ways in which DBRs can defend their position is by arguing that it explains intuitions in a wide range of cases (e.g. Kershnar 2000; Moore, 1993) and/or that it does not have unattractive implications. I will now challenge both these claims and the claim that the intuition is widely shared as we might first think. In doing so I aim to show that we have good reason to be wary of relying on the DBRI in justifications of punishment.

An alternative to the DBRI

I believe that there is a widely shared intuition that we should punish those who commit wrongs. Some DBRs interpret this as meaning that there is a widely shared intuition

that it is also non-instrumentally valuable for wrongdoers to suffer because they assume that this is the only explanation of the intuition to punish. But they are incorrect to do this. Recall that according to blame-based expressivism, though people are made worse off by having their citizenship rights removed, it is not non-instrumentally good that they suffer – withdrawing citizenship rights becomes permissible due to the changed relationship; and the positive reason to punish comes from the value of expression for its own sake as well as the value that expression has in maintaining relations with the victim. In this section, I aim to show that for many people the intuition that we should punish might be driven by the ‘*blame-based intuition*’ (*BBI*) that we have positive reason to express criticism of wrongdoers by negatively altering our treatment of them. Therefore, there is no need to appeal to any further idea that it is good for wrongdoers to suffer to explain this intuition. This challenges both the idea that the DBRI is as widely held as DBRs claim and that it explains our intuitions in as wide a range of cases.⁸

Stephen Kershnar and Michael Moore argue that the DBRI explains our intuitions in a wide range of cases. However, Kershnar and Moore have deterrence opponents in mind when they devise their test cases, so the cases they offer are focused on showing that many people still have the intuition to punish even when punishing brings no crime-reducing benefit. Kershnar, for instance, gives the following example:

Frank killed his spouse in order to cash in on her sizable life insurance policy. You, a prosecutor, know that this was a one time event and that he will not act violently again. Furthermore, you are certain that punishing Frank will have no net general deterrent effect. This you know on the basis of Frank's occupying a position of respect

⁸ I am not trying to argue that *no one* really holds the desert-based retributive intuition: I am certain that some do. I aim only to point out that not everyone who has a pro-punishment intuition believes that there is positive non-instrumental value in inflicting suffering on wrongdoers.

in a large crime-ridden community and his effective use of this position to get kids out of gangs and back into school. Through his actions, Frank prevents many incidents of gang-related violence that would otherwise occur. You are also confident that Frank's distinctive psychological makeup makes him immune from any moral reform that might accompany punishment. Through his speeches, Frank reinforces the moral fabric of a society and if he were convicted of first-degree murder there would be significant damage to this moral fabric (Kershnar, 2000: 104).

He assumes that we have the intuition that Frank should be punished. He then states, 'the case of Frank is generalizable' to cases where there are more crime-reducing benefits of not punishing and from this he concludes that the DBRI explains our intuitions in a wide range of cases (Kershnar, 2000: 104).

However, we know that as well as intentionally inflicting suffering, punishment is typically thought to have an expressive function and because Kershnar has deterrence opponents in mind, he has not controlled for this variable. Therefore, from Kershnar's example, we do not know whether we are dissatisfied with Frank's lack of punishment because no suffering was intentionally inflicted on him or because (we assume) there is no change in the way Frank is treated that expresses his changed moral relationship with other members of the community. Therefore, whilst cases like Frank might be sufficient to show we intuitively feel there is more value to punishment than deterrent value, they do not show that we intuitively feel that *inflicting suffering* is non-instrumentally valuable. To really test out whether this is our intuition, we need to consider a wider range of cases.

Let's consider a modified version of the Frank example.

Frank II: Frank killed his spouse in order to cash in on her sizable life insurance policy. You, a prosecutor, know that this was a one time event and that he will not act violently again. Furthermore, you are certain that punishing Frank will have no net deterrent effect. We decide to change our treatment of Frank but we do not aim to make him suffer. We remove his right to vote and restrict his right to move freely in the public spaces of the state. We do not give Frank the extra type of assistance usually reserved for co-citizens. Doing all this expresses Frank's changed relationship with other members of the community and that he has done something wrong. Also, let us stipulate that he does not suffer major psychological distress.

Frank suffers in the sense that his citizenship rights are removed. But what is our intuition here? Do we really think that Frank's suffering is positively and non-instrumentally valuable? I do not think so. I think for many people it is the fact that our changed behaviour expresses our changed relationship with and criticism of the Frank that it is important in explaining our intuition that we should punish him, not the fact that he is suffering. If Frank went on to take full opportunity of any educational or other opportunities when he was there so that his suffering was more or less counter balanced, would this really be a problem? I am inclined to think it is not. We have still criticised him in severe terms and excluded him from equal participation in our society. Not everyone will agree with me but this does not undermine the point: the fact that it initially seems intuitive that wrongdoers deserve punishment does not mean that we can assume that the vast majority of people think it has non-instrumental positive value for wrongdoers to suffer. It is plausible that the BBI is doing the work for a significant number of people; therefore the DBRI's position is weakened. If I am right then the DBR is neither as widely

held nor explains as wide a range of cases as the DBRs would have us think.⁹

Tadros has also provided an example that indirectly supports the claim that the BBI is doing the work in explaining the intuition that we should punish:

Hitler's Island: Although Hitler appeared dead at the end of WWII, he in fact escaped and lived the remainder of his life in isolation on a desert island. He never came to recognise that he had acted wrongly. Let's suppose that if Hitler's island was sunny, Hitler lived a happy life after WWII. If Hitler's island was rainy, Hitler lived an unhappy life (Tadros, 2013: 256-257).

Tadros has canvassed intuitive, and more considered, responses to the question of whether it would be better for Hitler to be on the rainy island and found them to be 'by no means uniform' (Tadros, 2013: 257). Tadros's example is useful in that it casts doubt on the idea that the intuition that it is valuable for wrongdoers to suffer is widely shared. It is also useful in supporting the idea that the BBI is doing a lot of explanatory work in cases like Frank. This is because there is no social context in the Hitler's island example, but there is a social context in the Frank example. In the absence of a social context it seems potentially acceptable for *nothing* (no suffering or another negative expressive response) to happen to the wrongdoer (Hitler). But, where there is a social context, it does not seem acceptable for *nothing* to happen to the wrongdoer (Frank). It would not be acceptable to carry on our relations with Frank as though nothing had changed. Similarly if we imagine that a passing cruise ship discovered Hitler and simply allowed him on board as an ordinary passenger, treating him in exactly the same way that any other passenger was treated, something would be amiss. This indicates that some *social* reaction such as the

⁹ One might wonder whether removing a wrongdoer's citizenship rights does not count as making him suffer.

blame-based expressivism (rather than the DBRI) is important in explaining the intuition that we should punish. This supports the idea that the BBI (rather than the DBRI) is doing the work in explaining why people have the intuition that we should punish.

Summary

In this section I aimed to cast doubt on the DBRI. I aimed to do this by showing that there are some cases, like Frank II, which suggest that the BBI (rather than the DBRI) might explain the intuition that we should punish wrongdoers for a lot of people. This undermines the reliability of the DBRI because it shows that it is not necessary to explain our intuitions in a wide range of cases and implies that the DBRI might not be as widely held as DBRs hope. I do not claim to have decisively dismissed the DBRI, but I do hope to have shown that philosophers should not take the DBRI as an incontrovertible fixed point in our moral reasoning. I think this, in conjunction with the fact that DBR recommends treating people in a way we usually consider deeply wrong, means that we should really avoid relying on the DBRI in our justification of punishment.

DBRs also face other kinds of objection for example: that they confuse what it is fitting for a person to have with what a person is *entitled* to have (Wellman, 2012; McDermott, 2004); and that they should be committed to the view that how ‘deserving’ a person is over the whole course of their life (rather than for the particular crime committed) is the relevant basis on which to punish her (Duff, 2001: 20; Tadros, 2011: 69-70). However,

these objections could be pressed on blame-based expressivism so I will discuss these in other parts of the thesis.

II. What's Wrong with Deterrence Theories?

As I said above, deterrence-based justifications of punishment claim that punishment is only permitted because of the instrumental benefits it brings in deterring crime: such benefits are required if punishment is to be permitted.

There are three ways in which imprisonment might reduce crime: incapacitation, specific or individual deterrence, and general deterrence. *Incapacitation* is the idea that imprisonment reduces crime because imprisoned offenders are removed from the general population and are thus physically prevented from taking part in criminal activity. *Specific deterrence* refers to the idea that the unpleasant experience of prison deters convicted offenders from committing further crimes once their sentence ends. *General deterrence* refers to the effect that punishment has in stopping people in the general population who would commit crimes were it not for the threat of punishment. The general deterrent effect of prison is generally agreed to be the most significant and it is this effect that most deterrence-based theories seek to justify, so my discussion in this section is focused on this type of deterrent effect (though I briefly explain the limited effect of the other two in the next section).

Pure consequentialism is the theory that only the consequences of an action matter to its justification. Pure consequentialist theories of the justification of punishment can be a type of deterrence theory if, as is usually the case, the consequence the theorist has in mind is crime reduction. However, I will not review purely consequentialist theories of punishment because pure consequentialism is a deeply unattractive moral theory. Suffice to say - as many have before me - that the consequences of actions are not the only important things in morality and it is difficult for consequentialist theories of punishment to *satisfyingly* explain why we should not punish the innocent if this will bring about good consequences (for criticisms see Hart, 1986; Duff, 2001: 7-8; Boonin, 2008: 41-50).

However, one does not have to be committed to *pure* consequentialism to be a deterrence-based theorist about punishment. One can believe that punishment is only justified if it serves the goal of crime prevention but also think that this is not the only relevant consideration to its justification (see Duff, 2001: 11; Hart: 1986; Tadros, 2011: chpt 2). In this subsection I review the principles underlying these kinds of positions. In the next section I raise concerns with the empirical *premise* that all these positions rely on.

One famous - though now quite dated - deterrence view is side-constrained consequentialism. This is the view that the goal of punishment is to reduce crime but that the pursuit of this goal must be limited by simple constraints, such as 'do not punish the innocent' (see e.g. Hart: 1986). However, this view suffers from the objection that it uses offenders as a means to serve the ends of others. This is because in so far as punishment is intended as a general deterrent, it is not intended to protect the general population from the offenders being punished but rather from those who would commit wrongs if there were

no penal institutions. Therefore it might reasonably be thought to use offenders as a means to protect the general population. This is a problem because we do not *normally* think it is permissible to use one person to benefit another, especially when the using takes the form of something we would normally consider to be wrongful (Duff 2001: 13-14; von Hirsch 1993: 9–14). It might well be that offenders are not used as a means (or at least not objectionably so) in the pursuit of general deterrence, but an explanation as to why this should be is needed.

One way of providing an explanation is by justifying general deterrence as a form of (or related to a form of) collective self-defence. Warren Quinn, Daniel Farrell and Victor Tadros have put views of this kind forward.

Quinn: In his famous paper ‘The Right to Threaten and the Right to Punish’ (1985), Warren Quinn argues that we can derive the right to punish from the right to threaten. Roughly speaking, the idea is that if it is justified to threaten people with punishment in order to prevent them from committing crimes, then it is justified to carry this threat out by punishing them if they commit a crime.¹⁰ We can formulate this as follows:

- (1) if at t_1 we are permitted to make a sincere threat that if x commits crime C we will give him punishment P , then
- (2) if x commits crime C at t_2 he ‘cannot object’ if we in fact subject him to punishment P (paraphrased from Quinn, 1985: 364).

¹⁰ A variant of this view is that the state is justified in punishing because it is a legitimate authority: to say that a state is a legitimate authority is just to say that it may coercively enforce the law, threatening punishment is assumed to be a way of coercively enforcing the law, so the state may punish.

The point is not supposed to be that they cannot complain because they have been forewarned, but rather I think it is supposed to be a logical point that they cannot complain at t_2 because they could not complain at t_1 .¹¹ If correct, Quinn's view would get around the 'using a means' problem because one does not use one's aggressor as a means if one harms her in self-defence. But these steps do not follow from each other: there is something new that X can complain about when we move from t_1 to t_2 .

Let's consider an example of a woman trying to deter a man from raping her. Let's assume that it is justified for her to sincerely threaten, 'If you rape me, I will kill you', that is, let's assume that she may make this threat and form the intention to carry it out. As I said above, in the self-defence literature, there is strong consensus on the idea that we cannot permissibly inflict defensive harm unless the necessity condition is met: we cannot permissibly inflict defensive harm if it is not needed to protect us. So it seems that at least part of the justification for the woman sincerely forming the intention to kill the rapist lies in the fact that it might protect her. And it seems that one of the major reasons the would-be rapist cannot complain is precisely that the threat is serving this purpose. If he complains that she is threatening him, she can say to him 'I'm just doing what I need to try and stop you raping me'. But after the rape has occurred, this is no longer the case. Killing the rapist is not needed for the woman's protection because the harm has already befallen

¹¹ The point cannot be that they cannot complain because they have been forewarned for this reason: Warning someone of a harm is not enough to make it permissible. We can see this using a variation on an example used by Jeff McMahan. He points out that if a police officer can reasonably foresee that his job is dangerous, it does not follow that it is permissible for criminals to attack him. The permissibility of the criminals attacking him would depend first and foremost on whether it was independently permissible for them to attack him (McMahan, 2009: 14). Likewise if the criminals warn the police officer that they are going to attack him, it doesn't follow that it is permissible for them to do so. It would only be permissible for them to do so if it were already independently permissible. But Quinn cannot assume that the threat we are warning of is independently permissible because this is the point we are trying to establish.

her. So after the rape has occurred there is something new therapist can complain about. He can say, 'Killing me won't stop the rape, it will not protect you. I have done something heinous but I am still a human being and it would be a violation of my rights and impermissible if you kill me'. 'The force of this objection still holds I think if we substitute the threat of killing with the threat of a more usual punishment. For example, two years in prison (which is the minimum sentence for statutory rape in the UK). The rapist can say 'Imprisoning me won't stop the rape, it will not protect you. I am still a human being and still have grounds to complain if you lock me up for two years.' Therefore, we cannot derive the right to punish from the right to threaten – to claim otherwise is to completely ignore the necessity condition that plays an important role in the justification of defensive violence.¹²

Someone might respond to this objection by pointing to Quinn's *Spike example*. He asks us to imagine that it is permissible for someone to defend their property by putting dangerous spikes on the outside of their garden wall in order to deter prospective burglars. These spikes do not physically prevent people from climbing over the wall, but they will harm the offenders as they climb. He thinks it is permissible to install the spikes and that any harm caused during the climbing is permissible harm (Quinn: 1985: 341-342). Then he asks us to imagine a situation where someone can only build the dangerous spikes on the inside of their wall. These spikes might deter a person from breaking in to his house but, if they do break in, the harm will only be inflicted on the burglar's way out i.e. after he's committed his wrong and when he no longer poses a threat. If it is permissible to build the spikes on the outside of the wall, then presumably it is permissible to build them on the

¹² I am not the first to put forward arguments of this kind (see also: Farrell, 1990; Tadros, 2011: 272).

inside. This, Quinn argues, supports the idea that we can derive the right to punish from the right to threaten.

But the example is crucially disanalogous to the way in which punishment works: once the wrong has been committed, there is a choice about whether to inflict punishment. In order for the spike example to be a true analogy, the house owner must have an option to deactivate the spike at the last minute. If he did have this option he should use it because spiking the intruder has no protective benefit so the necessity condition is not met. From the fact that it can be justified for a wrongdoer to suffer a penalty we have no choice but to inflict once the wrong has been committed, it not follow that it is justified for her to suffer this penalty if we have a choice about whether to inflict it. Therefore Quinn's spike example does not show that we can derive the right to punish from the right to threaten.

It may be true that *in general* threats need to be followed through in order to deter others from committing the same crime in the future – if you saw the house owner retract the spike at the last minute, then you would have less reason not to break into the house. But if we did not retract the spike in order to deter future offenders then this would need a justification because we are back again in a situation where we are using one person as a means to protect another (c.f. Tadros, 2011: 272).

Farrell: Daniel Farrell takes this standard claim from the literature on defensive violence: if a person is responsible for a threat to our safety, we may harm that person as a means of defending ourselves and by doing so we are not using them as a means in any

objectionable sense. He uses this claim as a way to try to motivate the idea that we do not use wrongdoers as a means when we imprison them in the name of general deterrence. He does this by arguing that they are responsible for future threats to our safety because, by setting a precedent of committing crimes, wrongdoers make it more likely for others to commit crimes (that is, unless we punish them). Therefore wrongdoers are liable to be punished because we are punishing them to prevent the future crimes that they are responsible for (Farrell, 1990).

This argument is able to avoid the charge of using offenders as a means. However, I do not find it very convincing. It implies that if I committed a rape, then although me having committed the rape is a necessary condition of me being punished, we would be not be punishing me *for* the rape but rather for my responsibility for future crimes. The initial wrong done by the wrongdoer seems to drop out of the picture. This is a problem for two reasons. First, it suffers from a ‘wrong wrong’ problem – if we are explaining the justification of my punishment to my victim, it is very unsatisfying to say that it is my responsibility for future crimes that renders me liable to punishment - surely it should be the wrong that I actually committed (c. f. Tadros, 2010: 276)?¹³ Second, an implication of Farrell’s account would be that the punishment received should be proportionate to the future crimes that others might commit rather than the crime I actually committed. Clearly a person who has a high degree of responsibility for X should be punished less than someone who has a low degree of responsibility for X. It seems to me that the degree of responsibility that wrongdoers can have for the future crimes of others is very small.

¹³ We may even wonder why it is necessary for me to have committed the crime at all. If we could show that I bear weak responsibility for a future crime, why is this not enough? However, (i) I’m not sure that would be a terrible problem in the sense that inciting racial hatred is in my view something that is a crime in virtue of the fact that it encourages future crimes and (ii) presumably Farrell would appeal to external considerations.

Therefore, it does not seem that Farrell's view would be consistent with the idea that the severity of a wrongdoer's punishment should match the severity of her crime.¹⁴ Therefore I do not think that Farrell's account is an attractive justification of punishment.

Tadros: Put simply, Tadros's *duty view* is that a wrongdoer incurs a duty to remedy the harm she has done to her victim. One-way of remedying the harm is to protect the victim and her associates from future harms. Wrongdoers may sometimes be required to bear harm in order to fulfil that duty. If a wrongdoer is punished, this helps to protect the victim and her associates from future harms because of the deterrent effect of having penal institutions in society. These duties are stringent and enforceable, so we can compel offenders to fulfil them even if they do not want to (summarised at Tadros, 2010: 17; 2012: 89-90).

Tadros's view avoids the objection that deterrent punishment uses wrongdoers as means because we are not (objectionably) used as means when we are made to perform a duty that we have. It avoids the 'wrong wrong' problem because the offender's duty to remedy is grounded on the wrong done to the victim (Tadros, 2010: 277). Some might find it odd to think of remedying criminal wrongs through preventing further harms. But Tadros makes the point that we cannot normally compensate people adequately for criminal wrongs like assault through monetary compensation and that the 'next best thing' to not having criminally wronged a person would be to help ensure they are not criminally

¹⁴ One could also read Farrell as arguing that wrongdoers are responsible for us being *less secure* against future threats, rather than for the threats themselves. However, I think even if that is the case, the arguments I apply above still hold. We would still be punishing them for the wrong wrong and they would still bear a fairly low degree of responsibility for our being less secure against future threats.

wronged again (Tadros, 2010: 277-278). Maybe a person could remedy through providing some other positive good – perhaps they have a special skill that they could use to provide something for the victim in a different way. However, I agree with Tadros that crime prevention seems appropriate and is an important goal that the state should pursue so, assuming punishment does prevent crime, I think that the duty can be discharged this way.

I find much of Tadros's view appealing and plausible. However, his view, like all the other collective self defence views, suffers from what I shall call the transferability problem. It also, like all deterrence views, is too reliant on the empirical idea that punishment is an effective way of preventing crime.

The Transferability Problem

All collective self-defence views rely on the idea that the harm one may cause in the context of self-defence is transferable into the context of punishment. This transferability is an important part of collective self-defence views: according to self-defence views, the reason we may harm wrongdoers in punishment is *because* we may harm them in the context of defence. Quinn's view, recall, is that if we are permitted to threaten a harm we are permitted to follow through on that harm. Thus, on Quinn's view, the harm we are permitted to threaten is *the same* harm we are permitted to inflict after the wrongdoing has occurred. In Tadros's view, we are permitted to harm a wrongdoer in order to protect her victim and associates from future harms in the context of defensive violence. He uses this idea to argue that we are permitted to harm wrongdoers through punishment because this is

a way of affording the victim and her associates such protection. Therefore, on the face of it at least, it would seem that whatever harm we are permitted to inflict on wrongdoers to protect the victim and her associates in the context of defensive violence, we would be permitted to inflict on her in punishment. This, I argue, results in Tadros's view permitting punishment that is inappropriate in type and Quinn's view permitting harm that is inappropriate in both magnitude and type.

In the context of defensive violence it is permissible to use physical harm to protect a victim. For example, it is permissible to threaten to kill or beat-up an attacker if this is the only way to stop him killing his victim. However, it is not permissible for the state to kill people or beat them up for attempted murder. But since according to Quinn we are permitted to carry out what we permissibly threaten, on his view, it would follow that this is permitted. Killing would be disproportionately harsh and thus it would follow that Quinn's theory can permit punishment that is inappropriate in magnitude. It also permits punishment that is inappropriate in type. This is because, since Quinn's theory would permit beating an attacker up in defence, it would also permit beating an attacker up as punishment. But it is barbaric and wholly inappropriate for the state to punish offenders through physical harm.

Tadros's view is different. He has the resources to put downwards pressure upon the magnitude of punishment. First, the idea that wrongdoers are liable to bear less harm to protect their victims from threats that they are not responsible for is intuitive. The main way punishment is supposed to protect victims from future harm is by providing a general deterrent to the rest of the population against committing crimes (i.e. harms that the initial

wrongdoer is not responsible for). So, according to the duty view, the wrongdoer would be liable to bear a lower magnitude of harm through punishment than she would be liable to bear to block the initial threat she posed to the victim (see Tadros, 2011: 349). Second, recall the necessity condition: in the context of self-defence we usually only think that it is impermissible to inflict a greater level of harm than is necessary to avert a threat. I argue below that we have good reason to believe that harsher sentences do not have an increased deterrence effect so this is likely to limit the magnitude of harm that it would be permissible for the duty view to justify in the name of punishment (c.f. Tadros, 2011: 351).

As Tadros himself argues this defence is limited: it is still left with the conclusion that, *in principle*, the duty view could permit a high magnitude of harm. For example, if it so happened that the instrumental benefits of the death penalty were very great then perhaps the duty view would not be able to rule it out (Tadros, 2011: 351). Also there are some cases where wrongdoers (perhaps very serious wrongdoers) may be liable to a high degree of harm to protect victims from future harms. However, on the whole I think the duty view can defend itself reasonably well against the transferability problem when it comes to explaining why it would not permit harm of an inappropriate magnitude.

But, I am less sure that it can defend against the transferability problem when it comes to considering whether it would permit harm of an inappropriate type. When we consider examples in the context of self-defence, we usually believe that we may protect victims by inflicting physical harm on their aggressors. Tadros believes that it is permissible to inflict physical harm on wrongdoers to protect their victims from further threats. For example, he thinks it is permissible to use them as a shield against a new attacker's gunfire (Tadros,

2011: 275). If this is permissible, it seems to follow that the duty view should license us to inflict physical harm on wrongdoers when we punish them. I think this implication highlights what I have come to believe is the foundational problem with collective self defence views of punishment. It may seem an obvious thing to say, but I have come to believe that defence and punishment are fundamentally different things. Defence protects us from harms (or wrongs) but punishment is our community's *response* to wrongdoing. Therefore, as we should expect, the types of actions that they permit are different. Whilst there are overlaps between punishment and defence and it is illuminating to look at one when trying to understand the other, there is a limit to the extent to which we can convincingly use a theory of self-defence to explain the permissibility of punishment.

I now consider some attempts to get around the transferability problem when it comes to harm type. One way would be to appeal to the necessity condition again. We could say that since long prison sentences do not seem to have a better deterrent effect than short ones, it may be a reasonable assumption that physical harm would not have a greater deterrent effect than the harm prisoners experience. If this is the case, we have no reason to assume that it would often be necessary to punish by physical harm so it would in fact only very rarely be permitted. However, this defence against the transferability of harm type would be limited in a similar way to how the defence against the transferability of harm magnitude is: it would mean that, in principle, the duty view could permit inappropriate types of harm (physical harm). I think the conclusion that physical harm may in principle be permitted is much more disturbing than the conclusion that an inappropriate magnitude of harm may be permitted. It is one thing to say that your view might, in principle, permit more of what we already think it should permit. It is quite another to say that it might, in principle, permit something else - especially when that something else is something we

normally think this a barbaric way of treating even wrongdoers. Also, the claim about longer sentences is that they are no more effective than short ones, not that they are less effective. Tadros's view can show why shorter sentences are to be preferred in these cases: we may not inflict more harm than necessary. But if it is the case that physical punishment is no more, rather than less, effective than other punishment, we need to find a reason why non-physical punishment is to be preferred to physical punishment.

Could a proponent of the duty view simply appeal that it is barbaric to punish wrongdoers using physical harm? That is, if it really is barbaric to punish using physical harm, then is it not open to the duty view to make use of this idea in explaining the impermissibility of physical punishment? If so, then the duty view could provide a principled way to oppose physical punishment. The trouble with this response is that we do not always think that physical harm is barbaric. There are a few contexts where this is the case: for example, we do not think it is barbaric to physically harm people when we do so to help them (as when we operate on them) and we do not think it is barbaric if people physically harm each other consensually (as boxers do). But most importantly for our purposes, we do not think it is barbaric to physically harm wrongdoers in defence of victims. If punishment is supposed to be a species of societal defence, then why should it be barbaric to use physical harm in punishment? If we want to explain why it is barbaric to physically harm people in punishment, I think it is best to admit that punishment is not the same as defence.

There is another way of trying to defend against the transferability problem when it comes to harm type. When trying to explain why his view can put downwards pressure upon the magnitude of permitted harm, Tadros appeals to the idea that the passage of time may

lessen it. He asks us to imagine that he could permissibly use you as a shield to avert a harm of magnitude x immediately after you have inflicted an initial harm on him. Then he says that if the threat he faces ‘occurs later in the day, or the next day, or the next year’ perhaps the amount of harm you are required to bear ‘diminishes accordingly’ (Tadros, 2011: 348). Perhaps he could say a similar thing about the type of harm: that though he may physically harm you to protect himself from a threat immediately after you have harmed him, he may not do so later in the day etc. The problem with this is as follows: although we sometimes think that the passage of time makes a difference to how we should or may respond to wrongdoers, it is usually only fairly long periods of time that shift our intuitions in this way and even then it is not clear that it is time itself that explains the shift. Perhaps the shift is explained by ideas such as the fact not many of the victims are left or that its significance for them has reduced; that the wrongdoer is now old, vulnerable, and weak; or that there is a large change in social context.

But for Tadros’s argument to work, it needs to be the case that an extremely short period of time is enough to make it appropriate to switch the way we respond to wrongdoers. If we were (by some miracle) able to convict and sentence a wrongdoer a day or a week after their crime, this amount of time would need to be able to explain why we were no longer permitted to use physical harm on the wrongdoer. But this is not in line with how time usually impacts our intuitions about wrongdoers: a day or a week would not normally be long enough to make any difference to our intuitions about mode or length of punishment. It seems to me more likely that time is not doing the work in explaining our intuitive shift. I think what’s doing the work is the fact that defence and punishment are different things and permit different kinds of harm.

A final attempt to get around the transferability problem when it comes to harm type might be this: we could say that, in the context of defensive violence, if there are two equally effective ways of protecting a victim from a threat that are equal in harm size, it is impermissible to use physical harm if we could use non-physical harm. Such an example is hard to construct because it is hard to say when physical and non-physical harms are of equal size but let's suppose that breaking someone's leg plus their recovery time from this is roughly equivalent to them spending three months locked in a moderately pleasant room with some, but not the usual level, of access to hobbies, friends, family etc. Now let's imagine that by some very strange twist of fate you have just been attacked by a culpable aggressor immediately after a different culpable aggressor has attacked you and you can either divert the second threat by breaking the first attacker's leg or putting him in the room. I confess I do not have a strong intuition here but the important point is that it is not at all obvious that it would be better to put him in the room. Maybe this is because this is a very unrealistic example where we have little other relevant experiences to help us form our intuitions, but maybe it is because defence and punishment are different things and permit different types of action: defence permits violence, but punishment does not.

In summary then I think that views which justify punishment as some form of collective self defence suffer from a problem when it comes to transferring the types and magnitudes of harm that we permissibly impose in defensive contexts to punitive contexts. I think that it is more difficult for such views to explain why the physical harm that typifies defensive violence should not be transferred to the punitive context than it is for them to explain why the greater magnitude of harm associated with defensive violence can be transferred. I think the transferability problem arises because defence and punishment are different things and so permit different forms of harm. I do not think this means that collective self-

defence views of punishment need to be abandoned or that my objection is decisive.

However, I *do* think it means that collective-self defence views should find a way to adapt so that they can take account of the different forms of permitted harm.

I think it is very important that we do not physically harm wrongdoers through punishment because it seems to go against the compelling idea that even wrongdoers do not forfeit their basic rights. However, perhaps I am wrong about this because perhaps there is no a significant moral difference between physically harming wrongdoers and harming them in the way that prison harms them. But this is nonetheless an issue that I think needs to be explored in depth by those who support collective self-defence justifications of punishment.

So far in this section I have reviewed some deterrence accounts of punishment and found most of them wanting at the level of principle. Now I shall move on to criticise the empirical premise on which they rely. I should note, however, that the next section is about the extent to which deterrence accounts can justify prison, rather than punishment in general.

Deterrence theory's empirical premise

We can prevent crime in different ways: by, for example, ensuring that vulnerable citizens have access to a decent range of opportunities so that it is easier to avoid turning to crime;

making sure that citizens are educated about the negative effects of crime on a criminal and society and the limited opportunities and harsh working conditions of many criminal careers; and deterring crime through attaching negative consequences to law breaking. There are also different negative consequences that the state can employ: for example, prisons of different categories, sentences of different lengths, fines, community service, and Home Detention Curfews (HDCs). So why does instrumentalism justify the use of prisons? One might suggest that the deterrence-based theorist only needs to adopt the claim that prisons deter crime and (assuming wrongdoers are liable to be made to bear costs towards this aim) that is a good enough reason to permit the state to use them.

But that is too quick. If punishment is only justified because of its instrumental value, then if two forms of punishment would create the same deterrent effect presumably it is wrong to use a harsher form.¹⁵ Just like expressive theorists, deterrence-based theorists need a non-arbitrary way of specifying why the state may use prisons. I suggest the deterrence-based theorist needs to endorse something like:

- (1) The *empirical premise* that there is reasonable certainty that prisons are necessary for (or one of the most cost effective ways of) preventing a significant amount of crime.

Deterrence views do not sufficiently question whether this is true and therefore whether their theories justify the use of prisons. Some of them do not even raise the issue (Quinn, 1985; Farrell 1990) whilst others raise the possibility but then just put it to one side (e.g. Hart, 2008: 2; Tadros, 2011: 41).

¹⁵ This is why I say ‘necessary for (or one of the most cost effective ways of)’ rather than just ‘is one way of in my formulation of the empirical premise.

However, if the empirical premise is not true then they need to supplement their theories with an account of why the state may use prison even if other forms of punishment would be less harsh. Consider, if the Duty View was not supplemented with such an account then if an imprisoned wrongdoer complained that he could equally well discharge his duty on Home Detention Curfew, there would be nothing the Duty View could say in response. In contrast, a blame-based expressivism has the resources to explain why it is permissible for the state to keep the wrongdoer in prison – the state is entitled to control its territory and so is entitled to insist that the wrongdoer stay in prison (so long as the wrongdoer has impaired her citizenship status to the extent that makes this response proportionate). In chapter 5, I suggest that blame-based expressivism would only permit the use of fairly liberal prisons, where inmates are given opportunities to spend time in their community and have frequent, informal and lengthy family visits. I also suggest and that there are some pragmatic reasons to sometimes use HDCs instead of prison. However, I think it is an important (though not decisive) advantage that the blame-based expressivist has the resources to permit the fairly widespread use of prisons. As I said in the introduction, it is desirable for a theory to be able to explain the permissibility of prisons both because the alternative is highly revisionary and because prison is an intuitive and attractive form of punishment.

It may be that deterrence theorists do *not* mind if their theories do not permit the widespread use of prisons. Tadros, indeed, even suggests that his view might ‘be a stepping stone to abolition’ (Tadros, 2010: 41). Nonetheless, it is still an important issue to explore. This is for three reasons. First, it is important for deterrence views to consider the

form of punishment that their theories justify because it affects the other details of their views. For example, when Tadros argues that we may enforce the duty to protect other citizens, the conditions under which this holds true must vary according to whether one is enforcing it through HDC or a US-style ‘supermax’ prison. Second, it is also important to know how radical your view actually is – if the widespread use of prisons is not permitted by your view this might affect your view of its plausibility. Third, is a strategic reason: as I just said blame-based expressivism only permits the widespread use of reasonably relaxed prisons. One might be tempted to reject my theory on that basis because one may be afraid that this would have an insufficient deterrent effect, but if actually we have reason to think that relaxed prisons are no less effective in this regard than more restrictive prisons, then one lacks this reason to reject my theory. I now move onto explain why I have concerns about deterrence theory’s empirical premise.

Worries about deterrence theory’s empirical premise

I said above that there are, three ways in which imprisonment might reduce crime: *incapacitation, specific or individual deterrence, and general deterrence*. I consider these in turn and conclude that none of them allow us to meet the empirical premise, or at least not to the extent that the widespread use of prisons would be justified.

Incapacitation: Incapacitation is the idea that imprisonment reduces crime because imprisoned offenders are removed from the general population and are thus physically prevented from taking part in criminal activity. The effectiveness of incapacitation may be

less than it first appears because crime rates are often extremely high inside prisons. There are also certain groups of offenders who have a low probability of reoffending so incapacitation is probably not necessary to stop them reoffending - and offending rates for those on HDC are typically low (Tadros, 2011: 30; Durlauf and Nagin, 2010 n.13). Nonetheless it is probably safe to say that incapacitation through imprisonment does prevent some crimes. However, it is worth noting that the number of offenders whom it is worth incapacitating is thought to be relatively small because only a small number of offenders commit large numbers of crimes (Durlauf and Nagin, 2010: 11; MacKenzie, 1997). Therefore any justification of prison based on incapacitation would only be relevant for a small number of offenders and thus would not permit the widespread use of prisons. In any case no deterrence theorist argue that the incapacitative deterrent effect of prison forms their main justification.

Specific deterrence: Specific deterrence refers to the idea that the unpleasant experience of prison deters offenders from committing further crimes once released. Though a significant number of ex-prisoners do not re-offend, it is fairly well established that prison is not particularly effective in achieving specific deterrence. Certainly no one is very happy with reoffending rates. and there is growing evidence to suggest that prisons are actually criminogenic: not only do they fail to deter many ex-prisoners from reoffending; they make them more likely to do so (Lösel, 2012: 990; Durlauf and Nagin, 2010). On reflection this is not that surprising: prisoners will be associating with other criminals which provides them with the opportunity to make contacts, buy further in to the criminal mentality, and learn new tactics; they are often largely cut off from their family support structure and other community ties which is thought of as a crucial factor in increasing

reoffending rates; and their employment opportunities after prison will be reduced due to the legal and social effects of having a criminal record.

This gives deterrence theorists reason to prefer other types of punishment to prison and suggests that the widespread use of prisons would not be justified by their specific deterrent effect.¹⁶

General deterrence: General deterrence refers to the idea that the existence of punishment deters people in the general population from committing crimes that they would commit if it were not for the existence of penal institutions. General deterrence is traditionally thought to have a large effect and deterrence theorists rely on this in their justification of prison. I am not going to argue that prison has no general deterrent effect. But can we be reasonably certain that the widespread use of prisons is needed to deter a significant amount of crime?

One methodological approach to the assessment of the general deterrence effect of imprisonment is to compare aggregate crime and incarceration rates between or within countries. Some studies of this nature show a significant negative association between imprisonments rates and crime rates, and have been interpreted a showing that prison deters crime. However, other studies claim that ‘there is no consistent correlation between prison numbers and levels of crime’ (Prison Refrom Trust, 2013). However, most

¹⁶ The poor reoffending rates amongst ex-prisoners may be because (British) prisons are overcrowded, underfunded, do not always give prisoners adequate opportunities to maintain contact with their family, and many prisoners leave prison without anywhere to live or work. Therefore these rates may not be a necessary outcome of using prisons *as such*. However, theorists are not typically trying to justify punishment on the basis of its specific deterrent effects so I put specific deterrence to one side.

criminologists question the validity of this general methodological approach because it is difficult to identify and control all the relevant variables (e.g. Durlauf and Nagin, 2010; Spelman, 2005). Durlauf and Nagin argue that one of the main reasons for this in the context of assessing the effect of incarceration on crime rates is that the causes of crime are incredibly complex and uncertain: 'there is no settled theory on the causes of crime let alone the appropriate way to quantify these causes' (Durlauf and Nagin, 2010: 14).

A different approach to examining the general deterrence effect of punishment is to look at the effects of specific (sentencing) policies on specific crimes in specific areas. This approach makes it much easier to control variables and so is thought to be preferable (though it does require a lot of data collection, some of which may be unreliable (Spelman, 2005)). Studies tend to look at the effects of the *severity* of, and *certainty* of punishment on the rates of specific crimes.

The severity of punishment is measured by sentence length. There seems to be a general agreement that increasing the length of a sentence has *no* statistically significant deterrent effect (Durlauf and Nagin, 2010: 14, von Hirsch et al, 1999; Ashworth and Roberts, 2012: 868; Doob and Webster, 2003). Of course as Doob and Webster say, no one is suggesting that 'a one-dollar fine for armed robbery would be the same as a three- year prison sentence. Rather, [just] that variation within the limits that are plausible in Western countries will not make a difference' (Doob and Webster, 2003: 191). This is important because it suggests that deterrence theorists cannot justify lengthy prison sentences on the grounds of general deterrence. But is also important to consider because if less severe prison sentences do not increase crime rates, would other less severe punishments that we

use follow suit? Would using HDC or community service orders make similarly little difference to the general deterrent effect of punishment?¹⁷ This suggests that we do not have reasonable certainty that the widespread use of prisons prevents a significant number of crimes.

The *certainty* of punishment is how likely an offender believes it is that she will be apprehended and convicted. There is a fairly high level of agreement that increased certainty of punishment has a meaningful deterrent effect (Durlauf and Nagin, 2010, von Hirsch et al, 1999; Ashworth and Roberts, 2012: 868). However, we should be cautious before assuming that this tells us very much about the general deterrent effect of punishment. This is because most of the studies focus on the role of the police in deterring crime (and they find that measures such as putting ‘more Bobbies on the beat’ and increasing the number of rapid response units significantly reduces crime) (Durlauf and Nagin, 2010: 23). Obviously, then, these studies only show us that increased police presence and the perceived threat of apprehension has a deterrent effect: they tell us very little about whether *prison* itself is needed. Perhaps an increased police presence and perceived threat of a different punishment, such as HDC or community service would be sufficient. So it is *plausible* to wonder about the necessity of prison, especially when we consider this point in conjunction with the evidence on severity (c.f. Durlauf and Nagin, 2011).

Clearly nothing I have said shows that prisons are not necessary for crime reduction (and I am sure some use of them is necessary). But I hope to have made a reasonable case that we

¹⁷ Unfortunately I could not find data on this, because the discussion tends to be focused on their specific deterrent effects.

have good reason to question whether the widespread use of prison is justified on deterrence grounds. Therefore we have reason to be wary about how much prison a deterrence theory can justify.

Conclusion

In this chapter I reviewed the many rival groups of theories to expressivism and rights forfeiture. I argued that the intuition on which desert-based retributivism depends should not be trusted because it recommends acting in ways that we usually think are wrongful; it is not needed to explain the intuition that we should punish and therefore it is not as widely shared as DBR's claim; and some people actively find the DBRI repulsive. Deterrence views rely on the empirical premise that there is reasonable certainty that prisons are necessary for preventing a significant amount of crime. However, we have reason to be sceptical about the truth of this claim at least to the extent that it can justify the widespread use of prison. Therefore, in so far as we take it as desirable for a theory of punishment to be able to explain the permissibility of reasonably widespread use of prisons; the deterrence theorists have reason to supplement their theories. This applies to deterrence theorists who draw on the ideas found in the justification of defensive violence. These collective self-defence views also rely on the idea that the harm that is permissible to inflict in defence is transferrable to the context of punishment. But this seems to leave them with the unwanted implication that inappropriate types of harm, such as physical harm, can permissibly be inflicted in the course of punishment.

I hope this chapter has shown that there are problems with both desert-based retributive and deterrence theories. Insofar as we want to justify punishment then, this gives us reason to turn to expressive and rights forfeiture views. However, whilst they have attractive elements, previous expressive and rights forfeiture theories suffer from problems. The task of the next chapter is to explain the virtues and vices of these previous theories, before going on to set out my own account in chapters 3 and 4.

CHAPTER 2

Previous Expressive and Rights Forfeiture Theories

Expressive theories of punishment seek to justify punishment on the basis that it is valuable or fitting for the state to express moral criticism of wrongdoers. Expressive theories are attractive. Unlike simple deterrence-based theories, they capture the idea that punishment is, and should be, about more than simply preventing crimes: it conveys a negative evaluation of the wrongdoer and it seems right that it should do so (see Feinberg, 1970). They are also preferable to desert-based retributive views in the sense that it is much more palatable and less controversial to think that wrongdoing merits criticism than it is to think that it is good for wrongdoers to suffer. However, previous expressive theories are plagued by what is often considered to be a fatal flaw: they cannot actually explain why moral criticism should or may permissibly take the form of the sort of hard treatment we would normally associate with punishment, such as, prison. Why is subjecting a person to hard treatment, by for example, sending her to prison, necessary to adequately convey criticism of their wrongdoing? The argument that expressive theories do not permit conventional punishment is not new and many theorists reject expressivism on its basis. In the next chapter I will argue that if we use Scanlonian blame as the relevant type of moral criticism, these problems can be overcome. But in this chapter I will explain and develop the argument that previous expressive theories do not justify prison or other types of hard treatment we associate with punishment.

As well as drawing on the expressive tradition of the justification of punishment, my theory draws on rights forfeiture views. Rights forfeiture views state that by committing a criminal wrong, a person forfeits some of the rights she holds and this is what permits the state to punish her (examples of rights forfeiture theorists are Wellman, 2012; Goldman 1979, 1982 and McDermott 2001). Rights forfeiture views start from an essential insight – in order for it to be permissible to punish someone, she must usually have forfeited her right against it (Wellman, 2012). However, they have been criticized as being unable to plausibly stipulate which rights an offender loses. Unlike expressive views, this is not because they are unable to specify that the state may respond to wrongdoing with hard treatment like expressive views. It is because the rights they are thought to specify as forfeited are not within the set of rights that would make them attractive accounts of punishment: Should rapists lose their right against being raped? Murderers murder? In the next chapter I will explain how my theory is able to avoid this objection, but in this chapter I will explain the objection more fully and consider other unsuccessful attempts to circumvent it.

In the first section of the chapter I explain why previous expressive theories fail to justify conventional punishment: their claim that conventional forms of punishment – by which I mean forms of punishment that use hard treatment, including but not limited to imprisonment - are needed to express criticism of wrongdoing seems arbitrary and they need to come up with a rationale as to why it is not. In the second section I consider and reject some influential attempts to argue that expressive theories do not fail in this way. In the third section, I discuss rights forfeiture views. My discussion here is shorter as there are fewer influential proponents, at least who put forward a positive account of which rights are forfeited by wrongdoers.

I. The Words Are Not Enough Strategy

Expressive theories of punishment rely on the expressivist argument, which consists of three main claims:

- (1) That the state should or may express criticism of wrongdoing.
- (2) That conventional punishment expresses criticism of wrongdoing.
- (3) That the state should or may express criticism of wrongdoing through conventional punishment.¹⁸

The primary purpose of this section is to see whether (3) follows from (1) and (2). In order to motivate support for the idea that it does, expressive theorists commonly appeal to the idea that mere verbal criticism of wrongdoing or doing nothing in the face of wrongdoing seems to be an inadequate response to it. Consider the following excerpts from expressive theorists:

[People] would surely see purely verbal condemnation of crime, however public and solemn, as half-hearted and unconvincing...it seems...that both those whose rights were being violated and those violating them, and everyone else for that matter, could not fail to draw the conclusion that those rights were...not really recognized...in any serious manner ...The necessary seriousness and weight can be secured only by [conventional] punishment (Primoratz, 1989: 200).

¹⁸ I take this breakdown from Nathan Hanna, though I have modified it slightly (Hanna, 2008: 128).

Just as calmly telling a friend she ought not to have lied to us communicates neither the pain she has caused nor our unqualified insistence that we not be so treated, so the state's verbal or written reprimand with attached explanation would be inadequate (Falls, 1987: 42-43).

I think the point expressed in the above quotes is absolutely right. If one does nothing in the face of serious wrongdoing or responds verbally but without affecting any further change in one's behaviour, this will seem unsatisfactory. Imagine that your sister had been bullying your child, and that you told your sister that what she did was wrong. Now imagine that, after this verbal criticism, you did not alter your behaviour towards her in any way: you did not reduce the extent of your interactions with her or withhold provision of any of the nice things you would usually do for or give to her. This indeed would seem unsatisfactory.¹⁹

The obvious problem for the above argument, however, is that it does not support the idea that (3) follows from (1) and (2). As others have pointed out, these philosophers draw on the idea that words are not enough to adequately convey criticism. But the fact that words aren't always enough does not mean that we must express criticism through conventional punishment: verbal criticism and conventional punishment hardly exhaust all the ways in which we can criticise wrongdoing (see Hanna, 2008; Sayre-McCord, 2001: 516-518). For example, as Igor Primoratz says,

We also express [moral criticism]²⁰ by adopting a cool, formal demeanour; by reducing the scope and intensity of our relations with the person concerned; by failing

19 Unless, perhaps, there are countervailing considerations, such as, that she has unstable mental health or you are worried about the impact of your actions on your nephews.

20 Primoratz uses the word 'condemnation' here. But if we are trying to assess expressive theories as a whole I think it best to stay neutral on the type of criticism under examination for the moment.

to render help or support in situations in which, were it not for the moral offence committed, these would be forthcoming without ado...All these are actions, not words; and they are used to express moral condemnation when mere words do not seem up to the task (Primoratz, 1989:198-199).²¹

These other non-verbal expressions of moral criticism are fairly widespread in our everyday applications of it. Thus from the fact that we should or may criticise wrongdoing it does not follow that we should or may choose to do this through conventional punishment. The claim that we must seems arbitrary.

At this point an expressivist could object that it all depends on how we understand the expressivist argument. If we understand it to mean that the state must express criticism of wrongdoing through conventional punishment because this is the only way of non-verbally expressing criticism of wrongdoing, then obviously the argument is false. This is because there are other ways of expressing criticism of wrongdoing. But, the objection goes, this is not how we should understand the expressivist argument: the idea is that the state cannot *adequately* express criticism of wrongdoing without using conventional punishment.

Expressive theorists often do appeal to this sort of idea and, though they do not usually spend much time clarifying it, the way it is presented suggests two understandings of this interpretation of the argument. The first understanding that expressive theorists hint at is that we risk incoherence or insincerity if we respond to serious wrongdoing with anything less than punishment. Consider John Tasioulas who claims that only ‘punishment enables us properly to *evinced the sincerity* of our condemnation’ (Tasioulas, 2006: 296 my emphasis).

²¹ Though he does then go on to ignore these other ways of expressing criticism when he makes his expressive argument for conventional punishment.

Nathan Hanna attributes this understanding of the expressivist argument to expressivists. But as he correctly points out, understanding it in this way is not particularly attractive because one ‘could react indifferently or even positively to a particular wrong or a particular wrongdoer without being insincere or confused about the relevant concepts’ (Hanna, 2008: 129). One might do this if, for example, one was dependent on the wrongdoer’s good will or if one wanted to try to reform the wrongdoer by showing her love and affection. The state might also then be sincere in its criticism of wrongdoers but refrain from punishing them because it will harm them or because it wants to reform them.

A better, and more charitable, way to understand the expressivist argument is as a *contextual or convention-based* claim: *given the norms of the society we live in*, we cannot adequately express the seriousness of wrongdoing without conventional punishment. To elaborate: in our society we could express criticism of serious wrongdoing with reactions other than punishment, but, given our conventions, this is inadequate. If we do not punish a wrongdoer we may express criticism of them without being insincere or incoherent. But, absent special considerations, we simply do not express adequate criticism; the criticism is not forceful enough. This is not a claim about the criticiser’s beliefs, it is about the contextual significance of her (in)action. In our society, not conventionally punishing wrongdoers has acquired an expressive meaning such that if an offender is not conventionally punished this signifies that the wrong has not been taken seriously enough (regardless of the beliefs of the would-be punisher). We can see shades of this in the above Primoratz quotation, when he says that without conventional punishment people

could ‘not fail to draw the conclusion that [the] rights [violated by the wrongdoer] were...not really recognized’ (Primoratz, 1989: 200, see also Tasioulas, 2006: 297).

Nathan Hanna, Matt Matravers, and Scanlon have argued that appealing to convention in this way is circular (Hanna, 2008: 131; see also Matravers, 2000: 78; Scanlon, 1986: 214). However I think this is too quick: the expressive meaning of our actions and words is necessarily shaped by conventions and context. Once an action or word acquires an expressive meaning through our conventions, we are, to a certain extent, stuck with it. Moreover, the meaning it acquires is real: if conventional punishment is required to adequately express criticism of wrongdoers in our social context then if we do not respond with conventional punishment, this really will be a case of failing to respond to the wrong adequately. This is not to say that we ought never to challenge contextually acquired expressive meanings, it’s just that when we do there are consequences to be considered. Thus we should be cautious before dismissing appeals to convention on the grounds of circularity.

However, I do not believe that *this* contextual understanding of the expressivist argument is successful in justifying conventional punishment. As I argued above, we have other conventional modes of expressing criticism of wrongdoing. And in other circumstances within our social context, these forms of criticism can seem adequate to express condemnation. Recall the example about my sister: if I were to withhold all friendly interactions with her and withhold help and support after learning she was bullying my child, this may well be an adequate response. I do not need intentionally to impose suffering on her too. Thus, if the state were to respond to wrongdoing with an analogous

form of behaviour, it would be staying broadly within the parameters of some of our conventional practices for adequately expressing criticism of wrongdoing. Therefore, claiming that conventional punishment is the only forceful enough way of the state's criticising wrongdoing seems exaggerated. Asserting that (3) follows from (1) and (2) seems arbitrary. Given the availability of these other options and the severity of conventional punishment, this claim needs an independent justification. Several influential expressive theorists have tried to do this by putting forward arguments that link expression and conventional punishment more firmly together. I will now examine these arguments and say why I think they are unsuccessful.

II. Other Attempts to Link Expressivism to Conventional Punishment

Antony Duff and Jean Hampton try to link expressivism with conventional punishment by giving more complex accounts of the purpose of expressive punishment. Primoratz and John Tasioulas supplement expressive considerations with the desert-based retributive idea that wrongdoers deserve that we inflict suffering on them. Thus we can only adequately express criticism of wrongdoers by inflicting suffering on them through a practice such as conventional punishment. I shall discuss Duff first, then Hampton. I challenged the type of desert-based retributive idea put forward by Primoratz and Tasioulas in the previous chapter, so I will not discuss it again here.

For Duff, the aim of punishment is to communicate to offenders the seriousness of what they have done, get them to realise this seriousness, and get them to repent (Duff, 2001: 28-30, 81-82).²² He believes that 'hard treatment' of the kind inflicted by conventional punishment is necessary to achieve these ends (Duff, 2001: chpt 3.4).

But is it? If we interpret the ultimate aim of Duff's theory to be to induce repentance in an offender then, as other theorists have pointed out, it seems the answer is 'no'. As these theorists argue, repentance cannot be thrust upon one from the outside: in order for repentance to be valuable, it needs to be genuine, that is, the wrongdoer actually needs to see why what she did was wrong and be sorry for it (see Hanna, 2008: 147; Matravers, 2000: 91). But, it does not seem likely that inflicting suffering on a person is a particularly effective way of getting her to repent:²³ it seems more likely that repentance will be achieved by coupling blame with intensive therapy rather than with prisons or, Duff's preferred option, community service (Hanna, 2008: 146; Narayan, 1993: 176-177).

It might now be objected that Duff's view is not that hard treatment is an effective way of securing repentance: it is that it is necessary for wrongdoers to undergo hard treatment if

²² Technically Duff's account is *communicative* rather than merely expressive. In expressive theories, the state simply aims to express its condemnation of the wrongdoer, in communicative theories the state aims for a deeper level of communication: it aims to get the wrongdoer to realise the seriousness of her conduct. The communicative element is important to Duff because he believes it is necessary in order to respect the wrongdoer as a responsible rational agent (Duff, 2001: 79-80).

²³ Certainly conventional punishment has high recidivism rates, as do programs that deliberately aim to make the victim's pain vivid (see e.g. Sayre-McCord, 2001: 515 n32)

they are to achieve true repentance, even if we cannot guarantee that it will achieve this (Duff, 2001: 108). But, as Deirdre Golash has argued, this is not an attractive position to adopt. Her argument takes the following form: if X is very valuable, but the only means we have of bringing it about is Y, we probably should not try to bring X about if we know that Y is certain to have a *pro tanto* negative impact and very unlikely to succeed. Therefore, even if hard treatment is a necessary means of bringing about (valuable) repentance we probably should not try to bring repentance about. This is because hard treatment is certain to have a *pro tanto* negative impact on offenders and is very unlikely to succeed. Thus (this interpretation of) Duff's argument does not succeed in supporting the idea that (3) follows from (1) and (2).

However, although Duff does argue that repentance is an aim of punishment, foregrounding this aspect of his argument is not the only way to interpret his work. In Duff's view, repentance is not the only aim of punishment. Punishment should also *reconcile* wrongdoers and their communities with each other by putting wrongdoers through a kind of secular *penance* (Duff, 2001: 109-111).²⁴ Penance achieves reconciliation because it constitutes a 'forceful and weighty kind of apology' (Duff, 2001: 109). The idea of penance is more successful than that of repentance in supporting the claim that the state may subject wrongdoers to hard treatment. First, we do not need to worry about whether conventional punishment is likely to succeed in bringing about penance: whilst the idea of forced repentance does not make sense, the idea of forced

²⁴ It might now seem that the communicative element of Duff's theory drops out of the picture. However, it does not. His view is still communicative because, whilst enforced penance is not the best way to induce repentance, it is still an attempt to communicate to wrongdoers the seriousness of their actions and it is still an attempt to get them to repent. It still treats them as moral agents because it still addresses them with good moral reasons that we could expect them to respond to (See Duff, 2001: 81-82, 89, c.f. Golash, 2005: 1189-119).

apology or forced penance is not as problematic: of course a genuine apology is better than a forced one, but the idea that subjecting someone to hard treatment can have similar reconciliatory effect as an apology is reasonably intuitive (consider ‘he’s served his time’, ‘she’s paid her dues’ etc.) and we see value in formalised public apologies even if we suspect they are not genuine (Duff, 2001: 109).

Second, it seems reasonable that penance does require some kind of suffering or burden to be borne by the penitent and therefore the idea of punishment as penance might plausibly put punishment within the realms of permissible actions. However, there is a problem with making (forced) penance the aim of Duff’s theory. If we interpret Duff’s aim as being to force offenders to recognise their wrongdoing and repent, it is plausible to think that this is good for them. It is plausible that their lives go better if their moral outlook improves and they register the appropriate response to their wrongdoing. However, if we interpret the aim of Duff’s view as forced penance, it is harder to work out why it is good.

Reconciliation is good, but I think that, although it is intuitive, the idea that imposing a burden on wrongdoers reconciles them with society depends on the idea that the burden suffered by the wrongdoer is purifying or that it cancels out the wrong done because it is somehow good that she suffers. This then starts to become mysterious (why would her suffering be purifying or good?) and starts to seem dangerously close to the desert-based retributivism. Duff distances himself from these kinds of ideas (Duff, 2001: 21-17) and in any case, I have explained why desert-based retributivism should be avoided in chapter 1. I therefore think that we should not interpret Duff’s work as claiming that the main justifying aim of punishment is forced penance. So Duff’s account still cannot explain why (3) follows from (1) and (2).

Even if one disagrees with my argument that Duff's view does not justify (3), there is a more external reason to be wary of it. Duff's view is very offender-orientated in the sense that Duff believes repentance and penance are good for and owed to the offender and both these points play a major part in explaining why we can justify punishment. I find the idea that genuine repentance is good for a wrongdoer attractive since it clearly increases virtue and can be understood as a sort of moral self-education. The idea that it is good for the offender to be restored to the community is also attractive. But I find the idea that we owe punishment to an offender or that it is justified because it is good for her very strange. Normally we can waive away what is owed to us. If this is the case, then the offender should be able waive away the punishment that is owed to her. But that obviously has very unattractive implications (c.f. Golash, 2005: 126).

Hampton's View

Jean Hampton argues that, when a person criminally wrongs his victim, he injures her by denying her value as an equal. She believes that when the state punishes the wrongdoer, it expresses that the wrongdoer is not above the victim and thus 'annuls the act of diminishment' (Hampton, 1992: 1687). She also refers to the idea that punishment restores the victim's sense of self-worth and makes her feel vindicated:

In short,[expressive punishment] is a response to a wrong that is intended to vindicate the value of the victim denied by the wrongdoer's action through the construction of an event that not only repudiates the action's message of superiority over the victim

but does so in a way that confirms them as equal by virtue of their humanity...To vindicate the victim, a...response must strive first to re-establish the acknowledgement of the victim's worth damaged by the wrongdoing, and second, to repair the damage done to the victim's ability to realise her value (Hampton, 1992: 1686).

If we understand the idea of 'vindicating the victim's value' as letting them know that we think they have equal value and that we care about them and that they have been wronged, then I think it is an attractive ideal that punishment of wrongdoers may achieve. However, Hampton is a little ambiguous about what 'vindicating the victim's value' means. She often seems to lean towards the idea that wrongdoing actually does damage the victims' value and therefore that punishment vindicates her value in the sense of actually restoring it. For example, she says:

[T]o be strung up, castrated and killed is to suffer severe *diminishment*...we are morally required to respond by trying to remake the world in a way that denies what the wrongdoer's events have tried to establish...thereby annulling the act of diminishment (Hampton, 1992: 1686-1687, her emphasis).

However I think the idea of vindicating the victim's value is less attractive when understood this way: though punishment obviously acknowledges the victim and the wrong done to her, being a victim does not in fact diminish a person's value so it is not obvious why punishment vindicates it in the sense of actually restoring it. In any case, whatever sense of 'vindicating the victim' we choose to use, it is not clear why we can only do this by conventional punishment and not by some other non-verbal form of moral criticism. I am also wary of Hampton's claim that punishment helps the victim regain her sense of self-worth: if *this* were our aim, wouldn't therapy or some other kind of support

for the victim a better option? Again it seems arbitrary to say that conventional punishment is the best way to proceed.

Summary

Thus far in this chapter I have argued that previous expressive theories do not justify (3): they are missing a non-arbitrary rationale that shows why moral criticism of serious wrongdoing must be expressed through conventional punishment. This is a seemingly devastating problem for such theories and some philosophers have rejected them on its basis (e.g. Hanna, 2008: 133-148; Hart 1963: 65–66). However, I think that (1) and the idea that words aren't enough to meaningfully express criticism of wrongdoing are intuitively appealing. Thus I think this seemingly devastating problem in fact raises an interesting question: what forms of negative treatment can expressive theories of punishment non-arbitrarily justify? I take this topic up in the next chapter, but now I will discuss previous rights forfeiture views.

III. Rights-Forfeiture Views

Rights forfeiture views state that by committing a criminal wrong, a person forfeits some of the rights she holds and this is what permits the state to punish her (e.g. Wellman, 2012; Goldman 1979, 1982; McDermott 2001; Morris, 1991). Blame-based expressivism therefore combines elements from expressive and rights forfeiture views because it argues

that wrongdoers temporarily lose their citizenship rights as a result of committing a crime. The idea that people can forfeit rights is not itself controversial. For example, it is not controversial that a culpable aggressor can forfeit her right against being killed if this is what needs to be done to save her victim's life. But rights forfeiture views of punishment have been widely criticized.

One of the main criticisms of rights forfeiture theories is that it seems difficult for rights forfeiture theorists to explain satisfactorily which rights wrongdoers forfeit. This is known as the problem of *suitability*.²⁵ Some have claimed that committing a wrong causes a wrongdoer to forfeit the same right (Morris, 1991: 70), or an 'equivalent' right (Goldman, 1979: 45), to the right that she violated. Let's call this *the strict reciprocity view*. But, as David Boonin argues this is very implausible – we would not be happy to say that a rapist forfeits his right against being raped and if this is the case then logically we should be unhappy with the claim that he has forfeited his right to something equivalent to rape (Boonin, 2008: 109-110, see also Lippke, 2001b: 79).

Many theorists have rejected rights forfeiture views on the basis of such criticisms (e.g. Boonin 2008: 109-110; Lippke, 2001b: 79). However, this is too quick a dismissal because rights forfeiture views start from an obvious but important insight: in order for it to be permissible to punish someone, she must usually have forfeited her right against it (Morris, 1991; Wellman, 2012; McDermott, 2001).²⁶ All theories of punishment must keep hold of this insight (but see Lippke 2001b). Therefore any plausible theory of punishment must

²⁵ Many other objections have also been pressed on rights forfeiture views, but I shall discuss them in chapter 5.

²⁶ I say 'usually' because it can sometimes be permissible to infringe a person's right against punishment if some further urgent good depends on it.

meet the dual challenge of honouring the insight from the rights forfeiture view whilst providing a plausible account of which rights are forfeited when a person commits a criminal wrong (c.f. Wellman, 2012: 378). I believe that blame-based expressivism can do this and that this is an important advantage of my view. I explain in detail why blame-based expressivism can do this later in the chapter. But for now I will explain why breaking free of the strict reciprocity reading of the rights forfeiture view is possible and consider another theorist's attempt to break free of it which I ultimately think is unsuccessful.

One might wonder whether a rights forfeiture view can avoid strict reciprocity. But why are philosophers so quick to ascribe the view to rights forfeiture theory in the first place? I can only find it explicitly defended in one of Goldman's papers (1979: 45) and mentioned in passing in Morris's work (1991: 70). I can see that there is an appealing symmetry to the strict reciprocity view which would make it seem like the natural form for a rights forfeiture view to take, but it is by no means logically required that it should take this form.²⁷ From the idea that a person forfeits some of her rights by committing a criminal wrong, nothing particular follows about *which* rights these are. The very specific idea that the wrongdoer forfeits the same right she violated (or its equivalent) is certainly not the only option open to the rights forfeiture theorist (c.f. Wellman, 2012: 386-387). We do not subscribe to strict reciprocity in other contexts such as rights forfeiture in defensive violence. If a culpable aggressor threatens to rape me, he forfeits his right against whatever harm is necessary and proportionate to prevent the rape – this may be death or lesser

²⁷ It might be logically required if you think the purpose of punishment is to 'cancel out' the wrong done. But (i) this is not clear (who is to say that only way to cancel the wrong is to 'pay' by forfeiting the same right? Two years in prison is not equivalent to rape but maybe this will do to cancel out the wrong?) and (ii) the idea that punishment 'cancels out' wrong is dubious and rights forfeiture theorists do not need to think that this is the aim of punishment.

physical harm but it is unlikely to be rape. So we need not assume that rights forfeiture views must be wedded to strict reciprocity – breaking free of this view is possible.

That said, however, though some philosophers have written to defend right forfeiture views from criticism (Wellman, 2012), to articulate the importance of their crucial insight (Wellman, 2012; Morris 1991), and to explain why rights are forfeited (Morris, 1991; McDermott, 2001), not many philosophers since Goldman really seem to have put forward an extended positive account of which particular rights are forfeited. Perhaps this is because we will never be able to give a fully determinate account of this, but it is important that a rights forfeiture view is able to say *something* about which rights are forfeited so that they can show how the suitability problem can be overcome. Blame-based expressivism does say something about this – it says that citizenship rights, but not basic rights, can be forfeited. Daniel McDermott (2001) has also put forward an account of which type of rights are forfeited, which could plausibly be thought to constrain the set in such a way that would overcome the suitability problem. Unfortunately, however, I do not think that McDermott's view is successful on its own terms in this endeavour and I also have some more external worries about it. I shall explain McDermott's view and how it might be used to combat the suitability problem before explaining my more external worries.²⁸

McDermott argues that when we wrong others we often harm them and/ or cause them to

²⁸ I should say that the main focus of McDermott's paper is on trying 'to demonstrate that there is a connection between wrongdoing and punishment such that the act of wrongdoing itself results in the specific changes in the wrongdoers' moral boundaries that make punishment [rather than simple restitution] morally permissible' (McDermott, 2001: 429). The suitability problem is only really considered in section V of his paper.

suffer some material loss but that the wrong is not reducible to just this harm or material loss. For example, he says that if Barry steals Arthur's car, Arthur suffers the material loss of the car, but this is not all. He also suffers a loss because Barry has violated his right that others not use his car without permission and therefore failed to treat him in a way that he is entitled to be treated (as a rights holder) (McDermott, 2001: 411). He believes that 'Barry incurs debts for value of both of these...[losses] - a debt for the value of... [the car], and a debt for the value of the treatment he withheld from Arthur by violating his right' (McDermott, 2001: 411).

McDermott labels the debt created by not treating another as rights holder as a *moral debt* and the value of treating another as a rights holder as a *moral good*; whereas the loss of the car is a *material debt* and the good of keeping the car is a *material good* (McDermott, 2001: 417, 415). Material and moral debts/goods, he argues, are fundamentally different in kind. According to McDermott, denying a person a *moral good* is denying him the *distinctive* good of being treated as a member of the moral community of persons.

McDermott claims that when we treat people in ways that are distinctive to them being members of this moral community, 'we treat them as 'one of us,' as if they are morally significant in a way that non-persons are not' (McDermott, 2001: 418). Examples of how we treat people in ways which are distinctive to them being members of the moral community includes being friends with them, marrying them, letting them build careers in our society as well as granting them rights. This is in contrast to material goods. We can give persons material goods but this is not, according to McDermott, distinctive to treating them as members of the moral community of persons because we can also give objects and animals material goods (McDermott, 2001: 415-418).

McDermott thinks the aim of punishment is to make wrongdoers 'repay' the debt they have created and thinks that moral debt, rather than the material debt, is the appropriate debt to focus on. Therefore he believes that we should punish wrongdoers by denying them something that is distinctive to being a member of the moral community:

By failing to provide their victims with the treatment they owe them as right-holders, wrongdoers incur debts to their victims for the value of this moral good, and, as a result, they forfeit their rights to other, equally valuable, moral goods. Punishment, according to this view, is a means of denying these forfeited moral goods to wrongdoers (McDermott, 2001: 424).

We can now see how McDermott's account might be able to overcome the suitability problem. When asked why we should not punish murders by murder or torturers by torture, he can say that protecting people from physical harm is a *material* good and therefore does not repay the *moral* debt created by failing to treat someone as a rights holder. With regards to torture, he says;

Torturing the wrongdoer in response to his crime would only be justified if the benefit of not being tortured were a moral good. And this, I believe, is simply not true. A moral good is a good the provision of which *necessarily* involves treating the recipient as a person. But the fact that I have not tortured another living creature is not sufficient for concluding that I have treated this creature as a person....I have never tortured my dog, but this does not mean that I have been providing him with the kind of treatment that distinguishes someone as a member of our moral community... if [the benefit of not being tortured] is not a moral good, then it follows that this good is not valuable in the same way as the moral treatment the torturer withheld from his victim,

and therefore the fact that the wrongdoer tortured his victim provides us with no justification at all for torturing the wronged (McDermott, 2001: 426-427).

As well as rebuffing the implication that the state should murder murderers and torture torturers, it would also prohibit the state from taking away wrongdoers' rights to goods like basic healthcare and food etc. as these are also material goods (McDermott, 2001: 426-427). If McDermott's account works, then this point is a huge advantage because it removes the connection with strict reciprocity and places limits on the cruelty of permissible punishment.

Moreover, McDermott also considers whether his account would justify the 'familiar, and intuitively appealing, form of punishment'; prison (McDermott, 2001: 428). He speculatively concludes that it might. This is because (at least in the prison McDermott envisages) prisons deny their inmates access to distinctive moral goods such as being part of our community, having a career in it, and participating properly in friendships and family relationships (McDermott, 2001: 428-430). As will become clear in chapter 5, I argue for a more liberal conception of what prisons should be like, but it would still be a considerable achievement for McDermott to be able to explain the permissibility of prisons in principle.

However, as I said, I have doubts about McDermott's theory. My first worry is that I am not convinced that it is successful on its own terms in circumventing the suitability problem. This is for two reasons. First, having a material good is not the same as having the *right* to a physical good. For example, one can lack a right against physical harm, and yet have the benefit of being free from it simply because no one has happened to harm

you. McDermott seems to have overlooked this distinction and this poses the following problem: although freedom from physical harm is a material good, the *right* against physical harm may reasonably be thought of as a moral good – on some conceptions of rights (including McDermott's) giving people rights is an example of how we treat persons in a way which is distinctive to them being members of the moral community. Therefore, although physically harming someone could not itself repay the moral debt generated by a wrong, removing the right against physical harm may do. So on McDermott's account, removing the same right that the wrongdoer forfeited might be an appropriate form of punishment. This would then put him in a position where he is vulnerable to the suitability problem: although there is no reason to think that his, or any other rights forfeiture view, is necessarily committed to the strict reciprocity view, he has not as yet succeeded in giving a positive guide as to which types of rights are forfeited. Thus, beyond intuition, he has no way of constraining the set of rights which may be forfeited and so has not fully broken free of the suitability problem.

Second, as we have seen, McDermott claims that wrongdoing typically generates two kinds of debt: a moral debt and a material debt. He appears to endorse the claim that material debts should be paid back (via restitution) as well as moral ones (via punishment) (McDermott, 2001: 413).²⁹ McDermott argues that Barry's moral debt should be paid off by denying him moral goods equivalent in value to the moral goods he denied to Alfred by violating his right. So, by the same reasoning, shouldn't Barry's material debt be paid off by denying him material goods equivalent in value to the material goods he denied to

²⁹ McDermott endorses the claim that material debt should be repaid in a brief comment rather than through detailed examination. But, I think it is reasonable to think he is committed to it. This is because he is consistent in his claim that depriving someone of a material good generates a material debt and saying that someone has a debt normally implies that you think they have something to pay off.

Alfred? If denying the moral goods to Barry is an appropriate response to the moral debt, then wouldn't denying material goods be an appropriate response to the material debt? This does not cause counterintuitive results of the sort relevant to the suitability problem when the material loss is the loss of property or money. For example, it is not a problem to claim that Barry ought to return Alfred's car (or monetary equivalent) to him as well as having moral goods denied to him to pay off the moral debt generated by violating his right. However, what of cases where the material loss suffered includes physical harm and pain? Imagine, for example, that Barry assaulted Alfred rather than that he stole his car. He would incur a debt to Alfred for the moral good he denied him by violating his right and also a material debt for the physical harm he caused. Unlike in cases where the material good is property or money this does generate counterintuitive results relevant to the suitability problem. If the material loss included is physical harm and pain, then shouldn't Barry pay off this debt by having this pain and physical harm inflicted on him?

McDermott could try to get around this problem in three ways. First he could say that physical harm does not count as a loss and therefore does not generate a debt. But that does not sound very plausible: surely my life goes less well if I have suffered significant physical harm, but why would that be if I suffer no loss? Second, he could say that physical harm does count as a loss but that this cannot be paid back by inflicting physical harm. I have sympathy with this view but it does not seem to be open to McDermott because if physical harm did not pay back physical harm then why would the loss of moral goods pay back the loss of moral goods? Third, he could say that the material debt could be paid off with money instead of by inflicting physical harm. However, I do not think that this option works. Although money can clearly make up for lost earnings caused by injury or the cost of treating it, it is not obvious that it is commensurable with actual physical

pain and any accompanying emotional trauma. Therefore giving money to victims would not *repay* the debt as such, it would merely be compensating them for their harm. This might be similar to how receiving a box of chocolates for a cancelled date does not repay you for the missed date, though it might compensate for it. It is perfectly acceptable, even good, to compensate people when we cannot pay them back directly. But, if we can pay them back in a more direct way, it is preferable that we do this. According to McDermott's account money cannot repay a moral debt (McDermott, 2001: 415). However, denying wrongdoers moral goods repays moral debts in a direct enough way to meet punishment's objectives. Therefore imposing physical harm should repay physical harm in a similarly direct way. Thus it would seem that paying back with physical harm would be a more direct way of paying back a debt caused by physical harm than money. Therefore, assuming that it is better to pay off debts in the most direct way we can, it seems that McDermott's account has the implication that physical harm would be better at paying off the a debt generated by physical harm than money. Thus I do not think McDermott can overcome the suitability problem – I think that his account implies that the best way to repay a material debt generated by physical harm would be to impose physical harm. I think this is a troubling implication (even if McDermott could come appeal to external considerations explaining why, all-things-considered, physical harm should be avoided).

My next worry about McDermott's account is that it doesn't really seem to explain why failing to treat someone as a member of the moral community of persons generates a moral debt that is repayable by the forfeiture of your rights to the equivalent moral goods. I think this is meant to be an intuitive claim (McDermott, 2001: 421). Of course we all have to rely on our intuitions at some point, but, as many have before me, I find this particular idea mysterious. It is unclear how denying him moral goods 'pays back' the moral goods lost

by the victim. Why does a rights withdrawal cancel out a rights violation rather than give us a rights violation plus a rights withdrawal? Is there a cosmic moral ledger on which rights violations are logged and must be balanced by rights withdrawals? What would be the point of such a thing? (c.f. Scanlon, 2008: 126). Blame-based expressivism is not mysterious in this way. It does not depend on any ideas about cancelling out rights violations. It depends on these common sense ideas: that the way we treat each other constitutes our special relationships with one another; that we can damage our special relationships with others if we wrong them; and that damaging our special relationships with others erodes the claims that arise out of these relationships. Moreover, even if we put the mysteriousness aside, the idea that punishing the wrongdoer pays back their wrong still rests on a problematic idea. It surely must rely on the idea that making wrongdoers suffer through punishing them is has positive non-instrumental value. If not, then, since you need a positive to cancel a negative, I do not see how it could offset the moral debt generated by the wrong. McDermott may well be perfectly happy with the idea that it is good for wrongdoers to suffer but, as we saw in the last chapter, there are many problems with it and I think it is best avoided.

Conclusion

In this chapter I discussed previous expressive and rights forfeiture views. As regards rights forfeiture views, my purpose was to: (i) to explain the importance of the insight from these views: any plausible theory of punishment must explain why the punishment does not violate the wrongdoer's rights; (ii) to describe the suitability problem and to show that existing rights forfeiture views have not succeeded in resolving it; and (iii) to examine

one particular attempt to do in greater depth. As regards previous expressive views my purpose was explain why they have found it difficult to explain why conventional punishment is in the permissible range of expressive responses to wrongdoing. Some people have rejected expressive and rights forfeiture views on the basis of the sorts of criticisms I have discussed in this chapter. However, I think both the crucial insight of the rights forfeiture view and the general sentiment behind expressive views (that is good for the state to express criticism of wrongdoing) are attractive. I do not want to abandon either of these ideas. Fortunately, I believe that blame-based expressivism can incorporate the crucial insight of rights forfeiture views whilst avoiding the suitability problem and stay true to the spirit of previous expressive views. I now move on to explain the detail of blame-based expressivism and how it can achieve this.

CHAPTER 3

Blame-Based Expressivism

If we want to know how to overcome the problems of previous expressive and rights forfeiture views, the place to start is by examining the notion of moral criticism in greater depth – what it means and what sort of behaviour it licences towards those who are legitimately subject to it. I argue that there is good reason to use Scanlonian blame as our preferred notion of moral criticism. According to Scanlon, when one person wrongs another, this impairs their relationship. He argues that, as well as making the judgment that a person is blameworthy for a wrongful action, blaming them is ‘to take that action to indicate something about the person that impairs one’s relationship with him or her, and to understand that relationship in a way that reflects this impairment’ (Scanlon, 2008: 123). The Scanlonian conception of blame as relationship impairment gives us a framework through which we can specify the ways in which the state may express criticism of wrongdoing citizens. If a relationship is impaired when one person wrongs another, then, absent overriding considerations, this means that special rights that were grounded in this relationship may be withdrawn or suspended. When one citizen criminally wrongs another this damages her relationship with her co-citizens and therefore I believe that the state may withhold citizenship rights from wrongdoers, that is, it may withhold some of the rights that its citizens have against it only in virtue of being its citizens. As I will argue in chapter 5, these include the right to vote and the right to move freely around the state’s territory. If the state can withhold the right to move freely around the state’s territory then we have the basis of an argument that the state can justify imprisonment. This argument

draws from expressive traditions and so can be seen as an expressive argument that can put prison in the range of permissible expressive responses to wrongdoing.

My argument also stays true to the crucial insight from rights forfeiture views. This because it explains why the wrongdoers have lost their rights against this treatment (they have damaged their relationships with the other citizens and therefore damaged the source of the special rights grounded in this relationship). It also breaks free of strict reciprocity because the rights lost are not the rights violated – the rights lost are citizenship rights. However, explaining why wrongdoers have lost their rights against punishment does not show that punishment is all things-considered-permissible: punishment costs a lot of money and we must explain why the state may spend that money on punishment rather than other social goods. I believe that the positive value of punishment comes mainly from the value of expressing criticism of wrongdoer. Expressing criticism of wrongdoers is valuable for its own sake, because it affirms the society's values and preserves a sense of solidarity amongst citizens and because, in cases where there is a clearly identifiable victim, it shows him that we care about the wrong he has suffered.

In the first section of the chapter I explain my decision to use Scanlonian blame, despite some of its pitfalls. Scanlonian blame gives expressive theorists a rationale according to which they can non-arbitrarily explain which rights can be suspended in the name of punishment (citizenship rights). In section II, I set out this rationale. In section III, I explain how blame-based expressivism can overcome the main problems with previous expressive and rights forfeiture views. In section V, I consider objections before concluding.

I. Why Use Scanlonian Blame?

Scanlon intends his account of blame as a comprehensive account of moral blame. But, as I said in the introduction, when conceived this way, his view faces objections. One of these is that we often blame those who we do not stand in any relationship with. For example we can blame Donald Trump for his irresponsible comments and actions even if we stand in no relationship with him. But, as George Sher (2013) has argued, it is not obvious how Scanlon's account of blame as a response to relationship impairment can account for this. Others have criticised Scanlon's conception of blame as not being passionate enough to adequately describe our blaming attitudes (Wallace, 2011; Mason, 2011). We often associate blame with attitudes such as anger, resentment or indignation (Strawson, 1981). But on Scanlon's account blame only requires the judgement that a person is blameworthy coupled with an attitude which reflects the idea that she has impaired her relationship with you. Therefore Scanlon's account does not obviously include some of our more passionate blaming responses.³⁰ I do not use Scanlon's account of blame as a comprehensive one. I only use it as an account of one kind of blaming behaviour that occurs inside relationships. But, even though I have qualified Scanlon's account in this way, we still might ask why I have chosen to use it as the basis of my theory of punishment.

³⁰ Though we should not that even though it does not obviously include these attitudes it does it does not explicitly exclude them either. Scanlon's account doesn't limit blame to one attitude. It therefore leaves room for the idea the idea that there are several different emotions that we might experience as we are blaming: one might feel resentment, rejection, or sadness and these can all be thought of as being part of our blame or following from it (see Smith, 2013: 38).

We can put aside the issue of blaming those who we stand in no relationship with for now. This is because, according to blame-based expressivism, we can think of state punishment as a form of blame between parties who are in a special relationship with each other: co-citizens.³¹ The main reason I have chosen to use Scanlon's conception of blame is that it relies on a series of straightforward and intuitive ideas – ideas which we can accept even if we do not think his account works as a comprehensive account of blame. The first is that our particular relationships with others are obviously partly constituted and can be changed by the current and past ways we behave towards each other. My relationship with my sister, for example, is partly constituted by the fact that we share parents and were brought up together, but our relationship is also constituted by the way in which we behave towards each other (how often we keep in touch, what we are in the habit of doing for each other etc.). Therefore, a relationship can be changed if the parties change the ways in which they behave towards each other. If my sister permanently emigrates to Australia, our relationship will change over time simply because we will not interact as much (c.f. Scanlon, 2008: 135). The second is that wrongdoing can impair our relationships. If our relationships can be altered by the ways in which we treat each other, then clearly they can be changed when one party behaves wrongfully to the other. For example, if my sister steals from me, this would change our relationship. Since it would change it for the worse, we can say our relationship is impaired. The third is that if I judge my sister blameworthy and adjust my attitudes and actions towards her in a way that reflects this, this is quite clearly an expression of moral criticism.

³¹ I will expand on the nature of this relationship, how it can be impaired by wrongdoing, and how this impairment can ground state punishment in the next chapter and in section II of this chapter.

In support of my claim that these ideas are straightforward and intuitive, compare the Primoratz quote I used earlier:

[We can express blame by] adopting a cool, formal demeanour; [and] by reducing the scope and intensity of our relations with the person concerned (Primoratz, 1989:198-199).

And:

Typically an act of wrongdoing brings about a distancing of the wrongdoer from the one he has harmed...and is felt as a breakdown or a distortion in the personal relations between the parties (North, 1987: 502-3).

So, even though Scanlon's account may have defects as a comprehensive account of blame, as an account of expressive blaming behaviour that occurs within special relationships, it relies on fairly attractive ideas. Since citizens stand in a special relationship with each other, this makes it good to use a basis for an expressive theory of punishment.

Another reason I have chosen to use Scanlon's account of blame is, perhaps surprisingly, one of the reasons it has been criticised: its lack of passion. As I said, Scanlon's account does not explicitly include some of our more vitriolic blaming attitudes. This might make it seem somewhat anaemic as comprehensive account of our blaming experiences, but it gives us reason to favour it as a basis for an expressive theory of state punishment. In chapter 1 I argued that one reason we should avoid desert-based retributivism is that the idea on which it relies – that wrongdoers deserve to suffer – is unduly harsh or barbaric. Similarly I think that, although all moral criticism must have some bite to it (or else it would not be criticism), it is preferable to base our justification of punishment on milder

forms of our blaming practices if possible. Therefore, since Scanlonian blame is milder in its starting point than some other conceptions of blame or other sorts of moral criticism, this provides a good reason to select it as our basis for an expressive theory of state punishment.

A final reason I am interested in using Scanlonian blame is purely a strategic one. As I explain in section II, Scanlon's work on blame can provide a framework through which we can non-arbitrarily specify the ways in which we may treat wrongdoers. Since this is what both expressive and right forfeiture theory needs, focussing on Scanlonian blame provides an exciting opportunity for someone wishing to draw on both these theories.

I have now explained why I have selected Scanlonian blame as the basis for my theory of punishment. Now I will consider its implications for how we may treat those who we can legitimately blame. In the next subsection I will consider the implications in the context of personal relationships; in the one after I will expand my discussion to the context of co-citizenship. However, first I would like to anticipate an objection. This is that Scanlonian blame is so mild that it cannot even be thought of as an example of moral criticism. R. Jay Wallace, for instance, has memorably accused Scanlon of 'leaving the blame out of blame' (Wallace, 2011: 350). If this were the case then it would be a problem for my account because, according to it, the main positive reason to punish is that punishment expresses moral criticism of wrongdoing. In response: there is a debate about whether the true meaning of the word blame necessarily captures elements of more vitriolic reactive attitudes such as resentment or opprobrium, or whether we can blame less harshly.

However, we do not need to get into this debate. This is because, whether or not one wants to call it blame, what Scanlonian claims is blame clearly *does* express criticism of wrongdoers. It is true that we can view our relationship with people as impaired without adopting a critical moral attitude towards them (as when, for example, a friendship gradually wanes). But Scanlonian blame is not *just* viewing a relationship as impaired – it also involves the judgement that a wrongdoer is blameworthy and that her blameworthy behaviour is what has caused the relationship impairment. When we view our relationship with someone as impaired as a response to their wrongdoing and this is coupled with the beliefs that she is morally blameworthy, it is quite clear that we are viewing them critically. When we make our actions reflect the relationship impairment in these cases, they are quite clearly an expression of this critical attitude. Thus if the state withdraws citizenship rights from wrongdoers whilst believing that they are morally blameworthy and that their blameworthy wrongdoing is what has made the withdrawal of citizenship rights appropriate, it is expressing moral criticism of them.

II. Blame-based Expressivism

What Does Scanlonian Blame Tell Us About How We May Treat Those We Are Permitted to Blame? (The case of special personal relationships)

Scanlon defines blame as the judgement that a person is blameworthy coupled with an attitude that reflects that your relationship with that person has been impaired by her wrongdoing. It seems safe to say that, where the blame is well placed, this attitude reflects a correct judgment: your relationship actually is impaired. Scanlon's primary concern is to

explain the conative element of blame, that is, his primary concern is to explain what *attitudes* it consists of, rather than what *behaviour* is appropriate. However, I think it reasonable to assume that, if an intention or attitude is appropriate then, all else equal, so too is a (proportionate) action that expresses it (and, indeed Scanlon himself makes this connection (Scanlon, 2013)). This equips us with the resources to identify appropriate changes in behaviour towards those we may blame using a clear rationale.

The rationale is as follows: being in a special relationship with someone gives you moral reasons to behave in certain ways towards her. If your relationship becomes impaired then the moral reasons to behave in these ways are weakened or removed. Therefore if we have standing to blame someone, then, absent special considerations, we may withhold or scale back the behaviours that we only had moral reason to perform because we were in a special relationship with her. Let me illustrate this using the example of friendship. If you are friends with a person, this gives you moral reasons to behave towards them in certain ways (such as, give them special assistance in their projects and put aside time to see them). If your friendship becomes impaired then this weakens the moral reasons you have to continue these behaviours. This is because the source of these moral reasons is the friendship and the friendship is damaged. A corollary of this is that relationship impairment does not give us a permission to withhold or scale back behaviours that we have moral reason to show to someone independent of our relationship with her (because we have moral reasons to show all people these behaviours). This is because if we have moral reason to behave in certain way towards a person that are independent from our relationship with her, the impairment of the relationship would not weaken the moral reason to perform the behaviour. For example, I have moral reason not to kill my friend. This reason is not grounded on the fact that we are friends (although our friendship may

give me extra moral reason not to kill you). Therefore if you wrong me, I still have moral reason not to kill you even though our friendship is weakened. This is because the source of my moral reason to kill you (not our friendship) is not damaged.

We can put this more formally:

- (2) Being in a special relationship to someone give us moral reasons to behave towards them in certain ways
- (3) If a relationship is impaired, then, *ipso facto*, the force of those moral reasons is weakened.

Thus

- (4) When a relationship is impaired it is permissible to scale back, withhold, or stop the behaviours usually arising out of the relationship.³² And
- (5) It is not permissible to scale back, withhold, or stop the behaviours that we ought to show that person that arise from factors other than our special relationship.

We will need to refer to independent considerations in determining what we owe a person in virtue of them being in a special relationship with us and what we simply owe to all people. But, what we have here is a way of helping to specify an appropriate range of behaviours for blame occurring within special relationships: we can suspend the behaviours that we have moral reason to show because of the relationship but not those that we have moral reason to show to people regardless of our relationship with them.

³² Of course, this idea applies in all cases of relationship breakdown, rather than just when the relationship breaks down because one party has behaved in a blameworthy manner.

We can put the rationale in terms of rights and duties: being in a special relationship with a person gives you special rights against her. If you wrong her, this impairs your relationship which means that the source of the special rights that you held against her is weakened. Therefore, depending on the extent to which the relationship is impaired, your special rights against her have been temporarily or permanently lost or weakened and it is therefore permissible for her to withhold the duties that correlated to the special rights that have been lost. Withdrawing rights that are owed to people simply in virtue of the fact that they are people and not in virtue of the fact that they stand in a special relationship with you is not permitted by the rationale. This is because the source of these rights has not been weakened.

A Challenge to Relationship Impairment

Before I move on to discuss the implications of this for expressive theories of punishment, I shall first consider an objection to the idea of relationship impairment. I have argued that wrongdoing impairs relationships and that this impairment causes the wrongdoers to lose (some of the) the rights that arose out of the relationship. I argued that this is a structural claim. The idea that wrongdoing impairs relationships and that this can cause wrongdoers to lose some of the rights that arose only because of the relationship seems pretty straightforward to me. But one might plausibly question it: Doesn't Scanlon's account misconstrue the nature of personal relationships? Aren't we meant to stand by those with whom we are in special relationships when they make mistakes?³³ Don't they retain rights

³³ I thank Alice Baderin for pushing me to consider this objection.

to remain within our special circle when they mess up? Should we really meet relationship impairment with more impairment?

I argue in chapter 3 that the state is permitted to withdraw citizenship rights from wrongdoers in punishment because co-citizens have the standing to blame on each other's behalf. Blame-based expressivist punishment is therefore mainly about whether specially connected third parties (i.e. the victim's co-citizens) may meet relationship impairment with impairment so this is the critical context in which I need to consider this objection. But it is also important to consider whether the victims themselves may meet impairment with impairment. This is because the rest of the citizens are blaming *on behalf* of the victim, so we do not want blame-based expressivism to rest on a prior implausible or unattractive account of how the victim may react when one party in a special personal relationship wrongs the other.

a. Blaming when you are the victim

There are some cases where withdrawing from someone who has wronged you seems impermissible. A classic example is if your child had wronged you in some way: as her parent, you have obligations to stick by her and help her not to do wrong. Angela Smith notes that, if, for example, the wrongdoer was your child, you might, depending on the context, have very strong reason to show love and affection to her, and to express criticism

to help the child learn to improve her behaviour in the future but not to withdraw from her (Smith, 2013: 39).³⁴

However, the existence of some cases where we ought not to manifest our blame does not show that blameworthy behaviour does not impair relationships or that, in the normal run of things, it is not permissible to manifest this impairment.

The idea that it is impermissible to meet relationship impairment with impairment becomes increasingly less plausible the more serious the wrong and the less close the relationship. Let's imagine a case of relatively serious wrongdoing where it is not plausible to think that the wrongdoer is excused (there are no mitigating circumstances and it is not the first time the wrongdoing has occurred). If the victim is not permitted to blame, then she is required not to blame. Let's think about what that means. A victim might well find not manifesting her blame humiliating, demeaning, or detrimental to her self-respect (c.f. Scanlon, 2013: 13). It seems reasonable that she should feel this way: if I did not modify my relationship with a cheating spouse, we might say, 'Where's her self-respect?' Surely (i) the fact that it seems reasonable for her to feel this way indicates that she would not be mistreating the wrongdoer in blaming him and (ii) we do not require victims to act in ways that undermine their self-respect (even though she may do this if she wishes).

One might reply that failing blaming need not always be demeaning. Often, if someone does not manifest blame to a wrongdoer, one might say it seems 'saintly' (Scanlon, 2013:

³⁴ Though she is discussing a case where the child has wringed a third party.

13).³⁵ I agree with this: *sometimes* it can be saintly. But as Scanlon notes, the use of words like ‘saintly’ confirms the fact that not showing blame is more than what’s owed. In other words, victims are not required not to, and thus are permitted to, manifest blame towards their wrongdoers.

Thus it seems to me that Scanlon’s view that wrongdoing impairs personal relationships is plausible and therefore that blame-based expressivism does not rest on an unattractive account of how the victim may react when one party in a special relationship wrongs the other.

a. Blaming when one’s special relation is the victim

We do not always have the standing to blame when we are not the victims of the wrong. But when the victim is someone who we are in a special personal relationship with, such as a close friend or family member, we often will have this standing. In chapter 3, I talk about some of the reasons why this is and argue that we have the standing to blame on behalf of our (victim) co-citizens. So now we need to ask, ‘In circumstances where you have the standing to blame on behalf of the victim, may you meet impairment with impairment?’

I believe you may. In fact, I believe you have positive expressive moral reason to. Though you have discretion over whether to blame on behalf of yourself, your discretion over

³⁵ Whether it is saintly or demeaning probably depends on one’s motivation for not blaming, the nature of the wrong, and the nature of the relationship.

whether to blame can be severely restricted when you are blaming on behalf of someone else.³⁶ I believe that you can sometimes be required, rather than just permitted, to blame when you are doing it on someone else's behalf. If someone seriously wrongs your child, for example, it does not seem you have discretion over whether to modify your behaviour towards the wrongdoer, it seems you *ought* to. The main reason that you ought to modify your behaviour is in order to protect the (valuable) relationship you have with your child. If your cousin had been systematically abusing your child, how could you not withdraw some of the duties ordinarily given to family members from him? In Michelle Mason's words, it would be 'pregnant with relationship-impairing meaning' for your child (Mason, 2011: 479). Not blaming would damage your relationship with your child, which provides weighty moral reason to manifest your blame of your cousin. Part of the reason that you may have reason to do this is that you want to let your child know that you care about her and that she has been wronged and that you think she has moral value. But it is worth noting that saying this does not rely (as Hampton's account sometimes seems to) on the notion that the victim's value has been diminished and therefore that it must be restored.

In this section I have argued that if we use a Scanlonian conception of blame this gives us rationale that can help specify the treatment we may give a person that we are in a special relationship with when they wrong us - it is *prima facie* permissible to suspend the special rights she had against you in virtue of your special relationship (that is, it does not violate her rights to treat her in this way). I said the correlate of this is that she may not suspend any rights she owes you that do not arise out of your special relationship because the force of these rights has not been weakened by the relationship impairment. I also claimed that

³⁶ I thank Ian Carroll for suggesting this.

we have positive expressive reason to suspend special rights when blaming on behalf of a third party.

What Does this Mean for Blame-Based Expressivism as a Theory of State Punishment?

Co-citizens stand in a special relationship with each other. This gives citizens special rights against each other.³⁷ These rights, *citizenship rights*, are grounded on the special relationship that exists between citizens, not merely on the fact that citizens are persons. Thus they are not owed to citizens of other states. They are also *not* rights that one would in principle hold against everyone but in practice hold primarily against your state because of pragmatic or other reasons (the right to protection from bodily harm may be an example of this type of right). They are rights that you *only* hold because of the special relationship that you have with your co-citizens. What these rights turn out to be depends on complex questions about what, if anything, makes the citizenship relationship special and what special rights, if any, citizens have over their territory. I am not going to discuss this now as it is reserved for chapter 5, but there I argue that citizenship rights are grounded on an ideal of equal membership which, amongst other things, gives citizens the right to vote and the right to move freely around the state's territory.

Obviously enough, individuals also have rights that are not grounded on the fact that they are a citizen of a country. They are simply grounded on the fact that they are persons. One example of such rights are *basic rights*. Basic rights are rights to what one needs to live a

³⁷ Citizens have these rights against states. This is true even though not all states accord all citizens these rights. All things being equal, if they do not accord these rights they violate their citizen's citizenship rights.

minimally decent life, such as rights against unnecessary harm. In chapter 5 I argue that basic rights also include rights to a minimum level of social inclusion, freedom to see loved ones, and opportunities to pursue activities that we value. However, there are other rights that are not basic rights – they are not needed for a minimally decent life - but are still grounded merely on the fact of our status as people, these might include the right to due process or the right not to be lied to (again I give this more explanation in chapter 5).

Above I argued that when our relationship with someone becomes impaired we may suspend the special rights that were grounded on that relationship. I also argued that when a citizen performs an act of criminal wrongdoing, this licenses all the citizens to view their relationship with her as impaired. Therefore, it follows that the special rights that arise out of this relationship, citizenship rights, can be suspended by the state when a citizen commits a criminal wrong. I also argued above that when our special relationship with someone is damaged, this does not give us the permission to suspend rights that we owe her which are not grounded on this special relationship. This is because the source of these rights has not been damaged. Therefore, when a citizen has damaged her relationship with her co-citizens, they may not suspend rights that are not grounded on the citizenship relationship (including her basic rights). This is because the source of these rights has not been damaged.

Above I argued that when someone wrongs a person with whom you are in a special personal relationship, this gives you positive expressive reason (as well as *prima facie* permission) to engage in blaming behaviour by suspending special rights she had against you. This is also the case in the context of citizenship. Of course, the co-citizen

relationship is not the same as personal relationships, but it is still valuable and ought to be protected. I contend that when one citizen criminally wrongs another, the state (acting as the representative of the rest of the co-citizens) must manifest its impaired relationship with the wrongdoer through action, or it would demonstrate a lack of regard for the rest of the co-citizens' positions as the victim's representative and for the solidarity they are supposed to share with her. This in turn would impair the relationship between state and victim. This means that the state has a positive reason to punish. (There is also other positive value in the state of expressing criticism of wrongdoing, both for its own sake and to maintain solidarity amongst the citizenry-at-large. But I expand upon this later).

There is an interesting implication here for crimes with a specific primary victim rather than crimes which primarily wrong all citizens. This is that, according to blame-based expressivism, we have a *pro tanto* reason to think that a victim has the power to waive the punishment of those who wrong her. If a victim decides for good reason that she does not want to express blame of a wrongdoer and that she does not want the state to express it on her behalf, then it seems like this may be up to her. Some people think this is counter-intuitive. I do not – I think there's actually something very attractive about the idea that victims should have this power. After all, *they* are the ones who've been wronged and, as we saw above, they might have good moral reasons for not wanting to express this blame. However, this does not mean this power cannot be overridden. First, even though it is very plausible that the victim has the power to decide not to express blame herself, it is less clear that she may demand those in special relationships with her to do the same. If your child grew up and forgave your cousin, I am not certain that you must do the same. Second, we should all be equal in the eyes of the law. Perhaps this is because it is unfair if some people that commit crime X are punished but some are not, so we may wish to

override a victim's power to waive punishment (if she has it) for this reason. Third, vulnerable individuals, such as those trapped in abusive domestic relationships may elect to waive the punishment when it would be gravely against their interests. So again, we may want to override the power to waive punishment.

III. How Blame-based Expressivism overcomes the problems of previous expressive and rights forfeiture views.

In the last chapter, I explained why previous expressive theories have not been able to provide a non-arbitrary rationale for why the state's moral criticism of wrongdoers may take the form of hard treatment, including typical forms of treatment that we would consider as punishment such as imprisonment. I also explained the challenge from rights forfeiture views. This challenge, recall, is that theories of punishment must take the crucial insight of rights forfeiture views seriously whilst overcoming the suitability problem. The crucial insight is any theory must be able to explain why wrongdoers lack rights against the punishments being inflicted. The suitability problem is that rights forfeiture views are thought to find it difficult to give a plausible account of which rights the wrongdoer forfeits. I made the ambitious promise that blame-based expressivism is able to explain non-arbitrarily why the state's moral criticism of wrongdoers may take the form of practices that we associate with punishment and also meet the challenge from rights forfeiture views. I am now in a position to explain how it does this.

Blame based-expressivism specifies how states may treat their wrongdoers: it may suspend their citizenship rights. As I said, in chapter 5 I argue that the right to vote and the right to move freely around a territory are citizenship rights. If the right to move freely around a state's territory is a citizenship right, then blame-based expressivism can start to explain the permissibility of prisons. Thus blame-based expressivism has the resources to specify at least two types of practices we associate with punishment. Therefore blame-based expressivism is able to non-arbitrarily explain why the state's moral criticism of wrongdoers may take the form of practices that we associate with punishment. If I am right, this is a significant step-forward for expressive theories.

Blame-based expressivism meets the challenge of right forfeiture views. My view is able to take the crucial insight of rights forfeiture views seriously as it offers an explanation as to why the wrongdoer does not have a right against what the state is doing. According to my view the state stands in a special relationship with its citizens. This relationship gives rise to citizens having special rights against the state (citizenship rights). When a citizen commits a crime, she undermines her relationship with the state.³⁸ Therefore, she weakens the grounds of her citizenship rights and thus some may be temporality suspended.

My view allows the rights forfeiture theorist to avoid the challenge that she must be committed to the implausible strict reciprocity view that permits the state to murder murders and rape rapists etc. As we have seen, according to my view the wrongdoing citizen undermines the source of her citizenship rights and thus it is these rights she

³⁸ One class of exceptions here might be civil disobedience and conscientious objection. I discuss these in chapter 5.

forfeits. She does not forfeit same right that she violated (as long as the relationship impairment does not undermine its ground). One might wonder whether the blame based expressivists should be logically committed to the claim that a citizen can only lose her citizenship rights by violating a citizenship right, such as the right against electoral fraud.³⁹ However, this is patently not the case. We can see this clearly by looking at the personal context. My friend can damage our friendship by violating rights that are not particular to friendship. She could obviously damage our relationship by assaulting me even though the right against assault is not a particular right of friendship. If I am right that the right to vote and the right to move freely in a territory are citizenship rights then blame-based expressivism has the resources to specify at least two types of punishment that are plausible in the sense that they fit more closely with our intuitions about the appropriate, or suitable, forms punishment should take. This therefore suggests that blame-based expressivism can take the insight of the rights forfeiture views seriously *and* overcome the suitability problem.

IV. Objections

Can Blame-Based Expressivist Punishment be All-Things-Considered Permissible?

In this chapter I have argued that blame-based expressivism can explain why wrongdoers lack a right against state punishment. But this only shows that punishment is *prima-facie*

³⁹ Of course all wrongdoers might violate a citizenship right because the right that others obey the law (as considered separately from the right that they not engage in harmful conduct towards you) might be a citizenship right. But as I argued above, making punishment a response to this wrong suffers from the wrong wrong and/or wrong victim problem.

permissible – it does not show that it is all-things-considered permissible. Maintaining a system of state punishment consumes an enormous amount of resources. These resources could be spent on other important goods, such as, healthcare or education (Tadros, 2011: 105). Punishment also imposes costs on the lives of wrongdoers and their families. These costs might be sufficient to mean that punishment is not permissible, even though it would not violate wrongdoer's rights. Therefore a theory that only explain why punishment is *prima facie* permissible is incomplete even though, as much of the vast literature on the justification of punishment demonstrates, it is quite an achievement to explain even just this. We also need to show why punishment has enough positive value to counterbalance these costs.

According to blame-based expressivism wrongdoers lose their rights against punishment because their relationship with the rest of the citizens has become impaired. The positive reasons to punish come from punishment's expressive value. But Victor Tadros has argued that the expressive value of punishment is not great enough to mean that the state may punish despite its costs – it is not enough to explain why punishment is all-things-considered permissible (Tadros, 2011: chpt 5). I disagree with this and I shall now explain why. I shall beginning by expanding a little on the expressive value of punishment. The first expressive reason for punishment is the value of the state expressing criticism of wrongdoing simply for its own sake. This value is an impersonal good. It is good that the state affirms the moral values of the community by criticising wrongdoers but it is not in itself good for anyone in any moral or material sense. Thus, whilst I think it is appropriate that the state criticise wrongdoers and that it does have some value, I do not think this value is very large and it is not where the main value of expressing criticising comes from.

The second positive expressive reason for punishment comes from the value that expressing criticism has in maintaining solidarity amongst citizens. When a person violates the law, there is a sense in which she wrongs all the rest of her co-citizens. Therefore, if the collective of co-citizens acting together through the state to express criticism of the wrong done to them all, this maintains solidarity amongst them. In cases where there is a clearly identifiable victim or small group of victims, the third positive reason for expressing criticism of wrongdoing is the value of maintaining the valuable relationship between the victim and the state (or collective of co-citizens). I think this is the most important positive reason for punishment that comes from blame-based expressivism. As I said above, if the state fails to make its interactions with the wrongdoer reflect the relationship impairment this is full of relationship impairing meaning for the victim. I also believe that by making its behaviour to the wrongdoer reflect the relationship impairment, the state also shows the victim that it cares about her and that she has been wronged. I think this aim is very valuable. We can see this by the fact that many people believe that the victim has a strong claim on the state that it punishes her wrongdoer. If the state does not, many people would believe that the state has let her down badly. Though there is no direct material gain to be had in the state maintaining its relationship with the victim and showing her that it cares about the fact she has been wronged, I think that it is something that there is strong reason to pursue.

Now we must consider the reasons against punishment and whether the positive reasons to punish can compete with them. The main ones I can think of are the impact on the wrongdoer, the impact on the wrongdoer's family, and the huge cost of running penal systems. Prisons under blame-based expressivism will have a negative impact on wrongdoers. However, I would like to remind the reader that blame-based expressivism

does not permit the removal of basic rights from wrongdoers. In chapter 5, I detail what a prison that respects the basic rights of wrongdoers might be like. I argue that such an institution must allow inmates extended, informal, and frequent family visits because we have a basic right not to be forcibly prevented from seeing our loved ones. Therefore the negative impact of prison on wrongdoers will not be as great as we intuitively believe based on our understanding of what many current prisons are like. Nonetheless, we need to consider whether this negative impact means that punishment is impermissible. I do not think it does. It might make a difference to whether withdrawing citizenship rights is the best thing to do but since we have already argued that having citizenship rights withdrawn is not a rights violation, it does not seem necessary to count the burden of having citizenship rights withdrawn again when we are considering all-things-considered permissibility.

The effect of punishment on wrongdoers' loved ones is more difficult. They have done nothing to lose their rights and it is likely that the imprisonment of the wrongdoer will have a negative impact on their lives. The first thing to say here is that this issue is a very difficult question for any theory of punishment. But according to blame-based expressivism, wrongdoers are entitled to frequent extended and informal family visits because this is a basic right. This might mean that it fares better than other theories. The second thing to say is that the wrongdoer's family should in principle get the same important expressive response from the state if they became a victim of crime. In general, they are likely to value this expressive response just as we all do, even though they suffer when their loved one is in prison. Thus, in so far as this is a valuable goal to pursue, it may be valuable enough to mean that their suffering is a necessary evil. We may also be able to

take steps as a society to try to mitigate their suffering, by for example, ensuring that they are not plunged into poverty as a result of their loved ones imprisonment.

The third issue to consider is whether the cost of maintaining punitive institutions makes them impermissible. The resources we use to fund our penal institutions could be used for other things, such as, healthcare Tadros makes the point that we generally value the harm reduction provided by deterrence more highly than expressive goals and that therefore only deterrence and not expressive considerations can have a significant role in explaining why the state may spend resources on punishment instead of things like healthcare (Tadros, 2011: 105). However, whilst I agree that on a case by case basis we do tend to prefer harm reduction to expressive considerations, I believe that expressive considerations can nonetheless explain why it is all-things-considered permissible for the state to pursue punishment.

We think our state should spend money on all sorts of non-urgent goods such as, sport, conservation, maintaining and opening parks, keeping our streets clean etc. We also think it should allocate resources to things that are more important than these non-urgent goods, but still not as important as harm prevention. Example might include educating people to a much higher level than they need to avoid harm or funding space exploration. We believe that it is worth taking a reduction in protection from disease and crime in order to live in a society that provides other goods. Of course, urgent goods like harm prevention and health care should take up a major chunk of the state's budget, and non-urgent goods, like maintaining parks, should probably not take up very much. But it does not follow that it is impermissible for the state to allocate significant levels of resources to goods that fall

somewhere in between urgent and non-urgent. Exactly how many resources it can devote to these goods will depend on things like how many resources the state has to start with and whether it has been able to provide a satisfactory level of urgent goods. But, assuming that our state is not extremely poor and that it is doing a reasonable job of providing urgent goods, it is permissible for it to spend reasonable quantities of resources on other things. I think the expressive value of punishment, especially the value it has in maintaining the relationship with the victim, is a really important good and I think that this is one of the things to which it is worth allocating a decent budget. Thus I think expressive considerations can play an important role in providing the positive reason to punish, not just a marginal add-on to the benefits brought by deterrence, and can therefore explain why punishment is all-things-considered permissible.

Can Blame-Based Expressivism Explain why Moral Criticism May Take the Form of Punishment?

Is my account vulnerable to the same criticism I levelled at previous expressive theories: that they seem unable to explain why moral criticism may take the form of punishment? Although blame-based expressivism explains why withdrawing citizenship rights does not violate citizenship rights, could we adequately express this criticism some other (less harmful) way? If so, then would expressive considerations actually support the use of punishment? Or could their aims be fulfilled by other forms of criticism? I have two responses. First, I do not think that blame-based expressivism is vulnerable to the objection I levelled at previous expressive theories. This is because my objection was that these theories are unable to put punishment in the permissible range of options through

which to express criticism – they do not even show that punishment is *prima facie* permissible. They do not provide an explanation of why wrongdoers lose their right against imprisonment. Blame-based expressivism does provide such an explanation and thus can at least explain why prison can be considered as falling within the permissible range and this is a significant advantage.

Second, above I accepted many expressivists' claim that, in the context of our society, action (rather than verbal condemnation) is needed to adequately express criticism of wrongdoing. But, I complained, we have actions other than punishment that we conventionally use to express criticism of wrongdoers. The examples of these that I gave were withholding the sorts of friendly and supportive behaviours which usually arise out of our special relationships with one other. I argued that previous expressive accounts do not show why the state cannot adequately express criticism by analogous types of action. However, my account explains why punishment actually is an analogous action. It is an example of removing the sorts of behaviours that usually arise out of our special relationships with others. Thus I could argue that, given our social context, action (rather than mere verbal condemnation) is required for adequate criticism and that punishment is an example of a way of expressing criticism that is engrained in our other social practices.

Does Blame-Based Expressivism Cherry Pick the Rights that can be Removed?

One might object that I making an unfounded: that the distinction between basic rights and citizenship rights maps on to the distinction between rights that can't be removed and those that can. However, whilst I do believe that these distinctions map onto each other, this not an unfounded assumption. It is no co-incidence that citizenship rights can be removed and basic rights cannot be. Basic rights are not grounded on a special relationship; therefore damage to a special relationship does not undermine them. Citizenship rights are grounded on a special relationship therefore damage to this relationship can undermine them.

It is true that we have to appeal to external principles to determine which rights are basic rights and which rights are citizenship rights. I do this in chapter 5. But it is important that we should retain a distinction between basic rights (and other non-relational rights) and rights that we only gain through our special relationship with our communities. On the one hand, basic rights are mean to possess a certain moral urgency and there are some things that we owe to others that do not seem to have anything to do with any special relationship in which we stand with them. On the other hand, our political communities are something special and I think it is important to recognize special rights that arise between their members.

I should be clear that I am not committed to the claim that people can never forfeit basic rights. One context where it is well established that they can is defensive violence. My

claim is that blame-based expressivism cannot explain why basic rights can be forfeited. Therefore blame-based expressivism cannot justify a system of punishment where offenders are deprived of what would normally be considered to be their basic rights. I think this is a virtue of my view. Many people believe that prisoners should retain basic rights (even though there is some disagreement about whether some rights should count, like the right to vote). My theory has an internal and natural way of explaining why this is.

Perhaps someone might pose the objection differently. They might say that there could be some deeper moral relationship that all people stand in and that basic rights are grounded in these. If this were the case then I would be making an incorrect assumption that the distinction between basic rights and citizenship rights maps on to the distinction between rights that can't be removed and those that can. Above, I talked about the idea that there is a 'default moral relationship' that all people stand in and that some behaviours (such as being friendly) might be conditional on that relationship being intact. However, the claim that basic rights are grounded in such a relationship is a highly controversial claim.

Normally we think that basic rights are grounded on a person's humanity or, as some prefer, personhood. This by its nature is not a relational property – one's having humanity or personhood is not dependent on one's relations with others. It remains for the proponent of such an idea to make the argument as to why basic rights might be relational in this way (though for something like it see Morris, 1991). And, as the wealth of writing on the justification of punishment shows, this is a very difficult argument to make.

Does Blame-Based Expressivism Confuse Fittingness and Entitlement?

Christopher Wellman has convincingly argued that what seems good or fitting does not affect what people are entitled to. Consider:

[I]magine that I want some of Bill Gates's money in order to fund release time from my teaching and administrative responsibilities, so that I can dedicate myself full-time to designing a new model of mosquito net that would better protect vulnerable people in sub-Saharan Africa from malaria. These admirable aims may well explain why I want some of Gates's money, but they cannot by themselves explain why I would be justified in taking it, because they say nothing about the various rights involved. In particular, absent some sort of explanation to the contrary, Gates stands in a privileged position of ownership over his money, and this position entails that I may not unilaterally help myself to it, no matter how laudable my aims are (Wellman, 2012: 372).

In this example what is good does not affect Bill Gates's entitlements. Wellman uses this example to argue against the desert-based retributive view that punishment is justified because it is good for wrongdoers to suffer. He convincingly argues that even if it is good for wrongdoers to suffer, this does not show that it is permissible to make them suffer because it does not show that the wrongdoer has lost their right or entitlement against being made to suffer. Blame-based expressivism does not claim that it is good for wrongdoers to suffer, but it does rely, in an indirect way, on the idea that it is fitting for wrongdoers to lose their citizenship rights. This is because according to blame-based expressivism, the wrongdoer is to some extent pushed out from the relationship of citizenship because that is thought to be fitting and it is as a result of this that she loses her citizenship rights. Does blame-based expressivism therefore fail to make the distinction

between fittingness and entitlements? If so, then blame based expressivism would not uphold the insight from rights-forfeiture theories.

In response: it might be true that some rights, like rights over one's money, are not affected by things like fittingness. But it does not follow that no rights –or 'claims' if you would prefer – are. The way in which one can come to lose a right surely depends on what type of right it is. Rights that are grounded on one's capacity for autonomous choice, for instance, can start to be lost as one loses this capacity in old age (even though one retains other rights based on one's interests); ownership rights over something are lost if you sell it; and, importantly for my purposes, special rights that one has on the basis of one's special relations can be lost if that relationship had been damaged or terminated. For example, let's agree that if you were on a boat trip with a friend and your partner, and you and your friend fell into the water and were at risk of drowning, you have a special right that your partner choose to save you over the friend. The source of this right is your special relationship with your partner. However, if you have done something to damage or terminate this relationship, the power of your right is weakened or perhaps even totally removed. For example, imagine you had damaged your relationship by abusing him or if your relationship had ended just because you split up several years ago and had not kept in touch. It seems to me that your special right for preferential treatment over your friend is considerably weaker or perhaps non-existent. This does depend on the idea that certain behaviours making it fitting to view your relationships in certain ways. But this cannot be thought of as a problem because it would be unduly limiting to say that our relationships cannot be changed by our behaviour. Whether the citizenship relationship is the sort of relationship that can be damaged by wrongdoing remains to be seen, but I make my arguments as to why it can be in chapter 4. For now, I think it is sufficient to note that one

should not immediately assume that blame-based expressivists confuse entitlements and fittingness.⁴⁰

Conclusion

In the last chapter I argued that previous expressive theories have not been successful in justifying conventional punishment. Expressive theories need a non-arbitrary way of specifying why the state may or must express criticism of wrongdoing through conventional punishment. In this chapter I argued that the way expressive theories can do this is by examining the notion of moral criticism. If we look at Scanlon's account of what it means to blame someone, we can see that when we wrong a person we impair our relationship with her and therefore any special claims that we had as a result of our relationship are weakened. The person we wronged would then, absent countervailing considerations, be permitted to suspend such special claims. In chapter 5 I will argue that one citizenship right is the right to move freely around the territory of a state so blame-based expressivism can permit prisons. Therefore blame-based expressivism is able to provide a non-arbitrary rationale which specifies why the state may or must express criticism of wrongdoing through conventional punishment. Blame-based expressivism also stays true to the rights forfeiture tradition of the justification of punishment. This is an advantage of the theory because any plausible theory of punishment should be able to explain why punishment does not usually infringe the rights of wrongdoers. Rights forfeiture theories have often been accused of not being able to explain which rights it is

⁴⁰ Some people might deny that our special relationship ground rights. Perhaps they are more comfortable with the idea that they ground special 'claims'. I think there is some plausibility in this idea so please feel free to substitute the phrase 'citizenship claims' for 'citizenship rights' in the thesis.

suitable for offenders to lose. The rationale I just mentioned means that blame-based expressivism is not vulnerable to this problem.

Blame-based expressivism can therefore be thought of as providing a step forward for both expressive and rights forfeiture theories because it is non-arbitrary and has the resources to justify prison. But aside from this, is it an attractive theory? I think it is: I think the thought that society should express criticism of wrongdoing is deeply attractive; I see it as important that a theory of punishment should have the resources to at least in principle permit the use of prison; and I think that Scanlonian blame provides an attractive account of what happens to our relationship with wrongdoers. However, there are a lot more questions to be answered because we can properly assess the attractiveness and plausibility of blame-based expressivism. For example, if blame-based expressivism is a reaction to a citizen's blameworthy action is it able to explain why the state may not punish things that we would normally consider private wrongs like adultery? Is co-citizenship really the type of relationship that can be impaired by wrongdoing? And how do we draw the line between basic rights and citizenship rights? In chapters 4, 5 and 6, I explain my theory more fully. In doing so I hope to answer these questions and others.

CHAPTER 4

The State's Standing To Blame

When a person does something blameworthy, we do not always have the standing to blame her. This can be for a number of reasons, perhaps, for example, we are guilty of the same fault or we were complicit in the wrongful act (see. Cohen, 2006: 118; Scanlon, 2008: 177). However, what interests me in this chapter is that we can also lack the standing to blame when the wrong done is simply none of our business (Smith, 2007: 478-479). If one citizen wrongs another then this is obviously the victim's business, it impairs the wrongdoer and victim's relationship and therefore the victim clearly has the standing to engage in blaming behaviour towards the wrongdoer. But if the victim is the only citizen who has the standing to blame the wrongdoer, then, at best, this could only provide justification for the state to permit or try to help the victim to suspend her own particular citizenship duties to her wrongdoer. Perhaps it could try to put in place mechanisms to prevent the victim's tax contributions from being distributed to the wrongdoer. But, as well as being terribly impractical, this would be insufficient because it would not be a communal response and, practically speaking, it would not be severe enough to meet our intuitions about what is considered to be proportionate punishment. If we want to ground a theory of punishment *by the state* rather than just by the victim, we need to show that the state may suspend all citizens' citizenship duties towards the wrongdoer. In order to show this we need to explain why all citizens have the standing to blame the wrongdoer. That is, we must explain why (i) the wrong done is the business of all citizens and (ii) that the wrong done impairs the wrongdoer's relationship with them as a group. This would

provide the state, acting as the collective agent of all the citizens, with the standing to blame the wrongdoer.

But we may wish to consider whether it is in fact an issue that legitimately concerns all citizens. In interpersonal morality, we often believe that people can lack the standing to blame when a particular wrong is none of their business. For example, if I am too merciless when I tease my brother, it is open to anyone to judge that I am blameworthy for it. However, if the paper boy starts engaging in blaming behaviour towards me for it we might think, 'What's it got to do with him?'. Most citizens are, in some senses, even further removed from most criminal wrongs done than the paper boy is from my treatment of my brother: most citizens are total strangers to each other and many criminal wrongs take place without other citizens even being aware of them. So why would all criminal wrongs be the legitimate concern of all citizens? For example, if one citizen steals from another, why would it be the legitimate concern of a third citizen who knows neither that the wrong has the wrongdoer, nor the victim?

The purpose of this chapter is to answer this puzzle. In the first section I review two unsuccessful solutions. In the second section I give my proposal: I start by identifying two features of a special relationship – representation and solidarity - which can be used to explain our licence to view the wrong done to another as our business in the personal context, then I apply this to the context of citizenship. In the third section I ask whether the co-citizenship relationship can be impaired: if co-citizenship were defined solely by sharing a place of birth or being subject to the same set of coercive institutions, then even if a wrong was all citizens' business, it would not alter the relationship between the

wrongdoer and her co-citizens. However, I argue that representation and solidarity are present in the co-citizenship relationship and can explain why co-citizens can view their relationship with a wrongdoer as impaired when he wrongs their co-citizen. Then I give a summary before considering objections and implications and concluding.

I. Unsuccessful solutions

Denying the Problem

One might wonder whether showing that all citizens have the standing to blame really does pose much of a problem: surely, one might think, everybody in the world can blame serious criminals (except perhaps other serious criminals). If everybody can blame serious criminals, then of course co-citizens have the standing to blame those who wrong their co-citizens. It is true that we can all judge all wrongdoers to be blameworthy and, in cases of the most serious wrongs such as murder it is probably not too controversial to think that most of us may vocalise this judgement, ask the wrongdoer to explain herself, or feel other disapproving reactive attitudes towards (c.f. Duff, 2010: 126). However, whilst these may be perfectly legitimate kinds of blame, blame-based expressivism relies on the idea of blame as a response to relationship impairment and it is not clear that everybody in the world may view his or her relationship with a wrongdoer as impaired. Consider the following: if you wrong me it is obvious that this action partly constitutes our relationship, and thus obvious that you wronging me can impair our relationship. But, but if you and I are UK citizens and you wrong a stranger who lives in the Australian outback, it is hard to

see how this defines our relationship. It does not therefore seem to follow that your action impairs our relationship.

There might be some wrongs which are so heinous that they can impair the wrongdoer's relationship with everybody. If you commit a truly heinous act, this might become a feature of your identity that no one can overlook. If this happens, then perhaps the wrong can impair your relationship with everyone. This is because it seems reasonable to say that a feature of a person's identity can partly define her relationship with others. For example, if you know that someone has suffered some very great tragedy, it might seem reasonable to think that this feeds into her identity and that this can shape her relationship with others. However, it only really seems intuitive to think that previous wrongdoings impair a person's relationship with everyone when we consider the most serious kinds of wrongs. For example, if I knew someone had assaulted another person unconnected with me when under the influence of alcohol and had expressed remorse I may need not consider this a defining feature of him or our relationship. But if I met someone who had systematically subjected his wife to mental and physical abuse over a number of years and had not expressed remorse, I would not be able to overlook this. I would see it as a defining feature of his identity. Thus in so far as we want the state to punish wrongdoing which falls short of being considered 'heinous' we have reason to avoid trying to justify punishment on the basis that everybody has the standing to view their relationship with wrongdoers as impaired.

There is another reason to avoid this approach. Scanlon argues that all people stand in a 'default moral relationship' with each other and that this can be impaired by blameworthy

wrongdoing. This idea has come under criticism – do we really stand in a relationship with someone we do not know and who has no other connection with us? (e.g. Sher, 2013: 50-55). However, let's set this to one side and accept that we do all stand in a default moral relationship. We now have to ask what sort of behaviour the impairment of this relationship would or would not permit: what sort of claims could plausibly be thought to be grounded on this default moral relationship that could be withdrawn when one is blameworthy? Scanlon does not provide any rationale for determining what behaviours are and are not grounded on the default moral relationship, but he is fairly clear that basic rights are not⁴¹ and he offers the following as examples of behaviours that are permitted in virtue of its breakdown:

[W]ithdrawal of trust[;] decreased readiness to enter into special relationships... with the person[; and] decreased willingness to help the person with his projects (Scanlon, 2013: 11, see also 2008: 143).

This, however, seems to be tantamount to saying that the sort of behaviour permitted by damage to the default moral relationship is unfriendliness (c.f. Mason 2011: 478-479). That may or may not be acceptable as an account of individuals engaging in blaming behaviour towards strangers, but it will not be enough to ground the kind of behaviour that we want the state to exhibit through punishment. I argued in chapters 2 and 3 that punishment can be explained as a version of our ordinary blaming behaviours in the sense that our ordinary blaming behaviours are often constituted by the withdrawal of special relational claims. However, punishment needs to be more than just collective unfriendliness. We therefore need to explain why more than just the default moral relationship is impaired. We cannot simply focus on the idea that everyone in the world

⁴¹ For example, he says that wrongdoers do not lose 'claims on us not to be hurt or killed, to be helped when they are in dire need, and to have us honor promises we have made to them' (Scanlon, 2008: 142).

can blame wrongdoers if we want to give a satisfying explanation of the state's standing to blame.

Law breaking Wrongs all Citizens

The next unsuccessful way of trying to explain the state's standing to blame centres on the idea that a citizen wrongs all citizens when she breaks the law. This would make the wrong the legitimate concern of all citizens and impair their relationship with the wrongdoer. One reason the wrongdoer might wrong all her co-citizens is that citizens owe it *to each other* to obey the law, so breaking the law wrongs all co-citizens. I think this is true. Therefore all citizens can usually view their relationship with criminal wrongdoers as impaired. This provides the explanation of why the state has the standing to blame when it comes to, *public wrongs*, that is, when it comes to crimes where all citizens seem to be more or less equally wronged. Examples of public wrongs might be speeding or tax evasion. However, when considering cases where there is a specific victim or set of victims, this style of explanation is insufficient because it suffers from a *wrong wrong* problem. In these kinds of cases, I do not want my state to punish the wrongdoer just because she has broken the law - I want it to punish her because of the other specific wrong she has committed. For example, I do not want my state to punish rapists just because they have broken the law; I also want it to punish them because they have committed an awful wrong (c.f. Duff and Marshall, 1998: 9). In these other kinds of cases it also and relatedly suffers from a *wrong victim* problem: I find the suggestion that the state only has the standing to punish rapists because rapists wrong all of us incomplete. On this account, if someone asked me 'what gives you the standing to blame the man who raped Susie?' the only answer would be

‘because he’s wronged *me*’. I think this does form part of the answer, but not all of it. We also want to be able relate our answer to the fact that Susie has been wronged. The wrong befell *her* and our response to it should not omit that fact. Thus basing punishment on the wrong done to all citizens would be unsatisfying as the only justification of state punishment – we need something else as well to explain our standing to blame for central cases of criminal wronging.

Duff and Marshall have put forward a different account of why all crimes are all citizen’s business and which therefore might be used as an account to explain the state’s standing to blame (Marshall and Duff, 1998; Duff 2010). They believe that the kind of act we should consider a crime is the kind of act where:

[T]he community should see the wrong done to [a victim] as a wrong done to ‘us’ [i.e. the whole community]’ (Marshall and Duff, 1998: 18). This is because the fact that we define ourselves as members of the same group means that ‘*an attack on the victim...is an attack on us*’ (Marshall and Duff, 1998: 19 my emphasis).

Perhaps then a Duffian might argue that all citizens have the standing to blame criminal wrongdoers because by wronging the victim they wrong us all.

This view wouldn’t suffer from the wrong wrong problem: it is the fact that the victim has been wronged that the fellow citizens are concerned about. They blame the wrongdoer *because he has wronged the victim*. But, this interpretation does suffer from the wrong victim problem. This is because our *standing* to blame him comes from the fact that, by wronging the victim, *he has wronged us*. Therefore if someone asked me ‘what gives you

the standing to blame the man who raped Susie?’ I would be answering, ‘Because he’s wronged *me*’ or ‘Because he attacked *me*.’

As will be evident from my discussion below, I think the idea that we are members of the same political community as our victim co-citizens is very important as to why we have the standing to blame their attackers. But, save in the case of public wrongs in order to explain the standing to blame we need to say something more than just that attackers wrong all citizens. We need an additional solution for why the state has the standing to blame in cases where there is an obvious primary victim, such as rape, murder, and assault.

II. My solution: Vicarious Blame

The idea that a person can impair her relationship with you by wronging someone who you are in a special relationship with is common in everyday morality. It is intuitive, for example, to think that your relationship with your husband would be impaired if he seriously wronged your child. In this section I will start by looking at personal morality: I will argue that if you are a representative of someone or share solidarity with her, this means that where someone wrongs her you are permitted to view your relationship with the wrongdoer as impaired. I call this ‘vicarious blame’ or ‘blaming on the victim’s behalf’. I will then apply my findings to the context of the co-citizen relationship and argue that, in this context too, solidarity and mutual representation, make it appropriate for one citizen to view her relationship with a wrongdoer as impaired when he wrongs her co-citizen.

Some Examples from the Context of Personal Relationships

It is not always be the case that we are permitted to blame those who wrong our special relations and, when it is the case, the reasons for it are likely to be different depending on the nature of the relationships involved. However, I want to point out just two different features of a relationship that can permit one party of a relationship to blame on behalf of the other: representation and solidarity.

I use teacher-pupil, team, and marital relationships as examples below. Of course, not all teacher-pupil, team, and marital relationships are the same: there is variation between particular relationships and between different cultures. But if I can identify features that typically exist in these relationships which licence vicarious blame, this will do for my purposes.

a. Representation

It seems to me that you can have standing to blame on behalf of those you are supposed to represent. I will use the example of the teacher pupil relationship to illustrate my point. In some respects, a teacher is meant to act as her pupils' *representative*: she will have to make decisions for her pupils and she is meant to take these decisions in their best interests. Thus, the pupil's interests become her legitimate concern. Of course, not all of the pupils' interests are the teacher's business, but at a bare minimum their educational interests are. So if something sets

back these educational interests, this is something of relevance to the teacher and where someone does something to seriously undermine these educational interests, the fact that this interest is relevant to the teacher can mean is permitted to view her relationship with the wrongdoer as impaired. Imagine, for example, that a teacher knows that the actions of a lazy colleague are going to undermine a pupil's educational interests. The colleague does not take time to assess student work thoroughly, plan lessons to meet the pupil's needs, or phone home to help implement behaviour that's conducive to learning. I think it very plausible that the teacher, in her role as her student's representative, can view her relationship with her colleague as impaired (at least in a professional context).⁴² Perhaps the teacher no longer needs to extend certain professional courtesies to the lazy colleague, such as trust in her judgment or reporting of a situation.

b. Solidarity

Another feature of a relationship that can permit vicarious blame is solidarity. Amongst other things, solidarity with someone in an activity is characterised by the view that you act together and, to some extent, are in it together, with regard to that activity. That is, you share the losses, benefits, and fate arising from the shared actions that are part of that activity (Pasternak, 2011: 9-10).

⁴² There might be very strong reasons for her not to manifest this impairment (as it might be unlikely to productively resolve the situation or it may be the role of the colleague's line manager to actually engage in a conversation with her). But this is a separate issue.

Loving marriages are characterised by solidarity across a very broad sphere: each partner should, for example, view a great many of the other's triumphs and losses as, to some extent, shared triumphs and losses which should be celebrated or endured together (Pasternak, 2011: 10). They should also be characterised by a sense of shared fate across a wide range of activities: if something significantly affects one of them, it significantly affects them both. It is fairly intuitive to think that this sense of shared fate can ground vicarious blame: if Sally wrongs Sarah's husband, then Sarah has the standing to view her relationship with Sally as impaired. I think this can be explained by the married couple's solidarity with each other.⁴³ The solidarity between them means that the wrong done is relevant to them both, is each other's legitimate concern: it is something they share in.

When I say, 'Sarah shares in the wrong her husband suffers' I do not mean that she is also a direct victim of the wrong. This would be implausible because she is not. If we say 'Sarah and her husband share in the praise he received at the awards ceremony' I might mean that they were both the objects of that praise, but I need not. I might mean that they share in it in the sense that them being in it together gives her a special moral reason to take special joy in it because, even though it happened to him, what happens to him is relevant to her because large parts of their lives are shared projects. In other words, their solidarity as a couple gives Sarah moral reason to view the praise as a legitimate concern of hers, and as something she has particular reason celebrate *together with him*. Likewise, saying 'Sarah shares in the wrong her husband suffers' can mean that her shared life and connection with him means that the wrong done to him is relevant to Sarah's life. It

⁴³ It may be explained by other features of the relationship too.

is something she has relevant reason to care about and is something she endures with him. If Sarah's solidarity with her husband makes wrongs done to him into something she shares with him, then it is plausible to think that this permits her to view her relationship with Sally as impaired.

Sharing solidarity with a person in one sphere of life does not automatically mean you share solidarity with her in other spheres of life. For example, the fact that I play on a football team with a person means I share losses, wins, and fate with her when it comes to our football team. But it doesn't mean I share in my teammate's successes, failures, and fate at work. Similarly, sharing solidarity with a person in one sphere does not permit me to blame on her behalf in other spheres: I might have a special permission to blame on my teammate's behalf if someone subjects her to a particular egregious foul, however unless we are also friends our team-based solidarity does not mean I can blame (in the sense of relationship impairment) on her behalf if she is wronged in another sphere. What matters in identifying the areas where I can blame on my teammate's behalf is identifying the areas we act together in: I can blame on someone's behalf when they are wronged in the sphere where we act together and share solidarity. In the case of the married couple, this sphere is large and nebulous; in the case of my teammate the sphere is quite sharply delineated.

The Standing to Blame amongst Co-citizens

I think that representation and solidarity are features of the relationships between citizens in a democracy and that these features of that relationship can justify citizens' standing to blame on each other's behalf. I now explain why.

a. Representation

It is fairly uncontroversial to assume that most citizens in a democracy believe the state is or should be their representative: the decision making body used by the collective will take decisions on the rest of the citizens' behalf and it is supposed to make decisions in their best interests. Not all of a citizen's interests need to be taken into account by the decision making body, but those that do that do are the state's business. Not being criminally wronged is a fundamental interest of citizens and it is fairly uncontroversial to think that this interest is one which is the business of the state. A state must have a system of criminal law and most citizens in a democracy believe that its violation is the state's legitimate concern. Thus if a citizen is criminally wronged, this sets back an interest that is the state's legitimate concern. Since the state is just an agent acting on behalf of all citizens, this more or less amounts to saying that it is *all* citizens' legitimate concern when one of them is criminally wronged. Thus when one citizen criminally wrongs another, this is relevant to all citizens' relationship with the wrongdoer and permits them to view this

relationship as being impaired.⁴⁴ Therefore the state, acting on their behalf, is permitted to withdraw the wrongdoer's citizenship rights.

b. Solidarity

Here I draw on Anna Stiltz's work in *Liberal Loyalty*. Stiltz argues that people have a duty to form democracies. I accept this as an uncontroversial premise. She also argues that democracies are collective agents, that is, members of a democracy act together as a group in certain key processes (Stiltz, 2009: 206). Stiltz makes use of Michael Bratman's model of collective action. Put simply, this model states that we must be acting with (individually held but) shared we-intentions in order to act together (that is, act collectively). A 'we-intention' is simply the intention to do something *together* with others: it must 'reference the fact that [the individuals involved] are doing *something as a group*' (Stiltz, 2009: 192-194, my emphasis). All members must also know that the others share their we-intention in order for collective action to be achieved. For example, if I want to go to a party, I must go somewhere where there are other people. But if my intentions are entirely self-referential, say I am just going to have a good time, there is an important sense in which my fellow partygoers and I are not *acting together* as a group. By contrast, if Sarah and I throw a party and we are both intending to throw a party together and both know each other's intentions, then we clearly act together.

⁴⁴ Notice this is not the same as saying that all citizens are wronged by the wrongdoer, only that the wrong done is relevant to all citizens.

Stiltz argues that members of a democracy share we-intentions. We *can* perform many actions associated with democratic citizenship without a shared intention, such as, paying taxes or obeying the law for purely prudential reasons (Stiltz, 2009: 205). However, Stiltz argues, a citizen of a democracy ‘has the obligation to deliberately involve herself in the political process’ (Stiltz, 2009: 205). Inspired by Rousseau, she believes that participating in the political process (morally) requires one ‘to be capable of taking up the viewpoint of the general interest or common good’ (Stiltz, 2009: 35). If a citizen does this, then she has a we-intention. Stiltz gives some examples of when citizens act with a we-intention. These examples include debating political issues, joining social movements, voting, and retaining a right of participation in revisions to the constitution (Stiltz, 2009: 206-207). I endorse Stiltz’s view that citizens ought to view the democratic process as one where citizens give each other’s interests moral weight, rather than as a *modus vivendi*. But I do not believe that that we need to take such an active view of what is required as Stiltz does. It seems to me that insofar as citizens of a democracy view themselves as living in a political community which is, in some sense, being run by all the members together, they in some sense share a ‘we-intention’.

This is relevant to whether citizens have the standing to blame on behalf of the victim because some of the we-intentions in the democratic context creates solidarity. Running the state is a shared activity and, since democracy is more than a *modus vivendi*, citizens are ‘in it together’. They share in the success and failures of its endeavours. As I said above, sharing solidarity with a person in one context doesn’t mean you share it with her in all contexts: what matters is whether it is an area where you act together. But if citizens should have we-intentions about anything, it seems minimally controversial that they should have we-intentions about establishing, maintaining, and obeying a criminal law.

Thus it seems to me that if a citizen is criminally wronged, she suffers in an area that forms part of citizens' collective actions. She suffers in an area where citizens have solidarity..

Thus it seems to me that citizens' relationship of solidarity with the victim can permit them to blame the wrongdoer. Just as Sarah's solidarity with her husband and my solidarity with my football teammate makes the wrongs done to them specially relevant to Sarah and me, citizens' solidarity with the victim in the area of the criminal law makes criminal wrongs done to him specially relevant to them. Thus I believe that citizens' relationship of solidarity with the victim can permit them to view their relationship with the wrongdoer as impaired.

Thus far in the chapter I have argued that victims have the standing to blame their attacker and I have argued that, because of solidarity and representation, the citizenry at large has the standing to blame on behalf of the victim. This is an important step in explaining why the state may punish wrongdoers, however I still need to explore the idea that the co-citizen relationship is one that can be impaired. I will move onto this shortly, but first it is worth mentioning a few of favourable implications of blame-based expressivism that arise out of the arguments just given.

My view does not suffer from either a wrong wrong problem or a wrong victim problem. It does not suffer from a wrong wrong problem because it is the fact that a wrongdoer has wronged *the victim* inside an area where we share solidarity with him, rather than the fact he has broken the law *just as such*, that causes the problem for our co-citizens. It does not suffer from a wrong wrong problem because on my view, the answer to the question 'What gives you the standing to blame Susie's rapist?' is 'Because my relationship with the Susie

means that the wrong done to her is of special relevance to me' not 'Because the wrongdoers has wronged me by wronging Susie'.

Blame-based expressivism also deals adequately with cases where we lack a specific victim, such as cases of speeding and tax evasion. As I said above, it when a citizen breaks the law it is plausible to think that he has have wronged us all and it is properly all our concern so this gives us all the standing to blame. In these cases, citizens are blaming on behalf of themselves and this is one significant way of deriving the standing to blame in these cases. However in such cases citizens are also still blaming on behalf of each other. They are all wronged by the wrongdoing, and the wrong that befalls each citizen is the legitimate concern of all other citizens. Therefore, when they blame they may do so on their own behalf but also on behalf of the other citizens. When they blame on behalf of each other, they are also maintaining their solidarity – they have all been wronged and, because they are in it together, this is something that they should respond to together and for each others' sake as well as their own. This is important in my account because one of the main positive reasons to punish is that it maintains our relationship with the victim. Therefore if citizens were only blaming on their own behalf in cases of public wrongs then this positive reason to punish might be lacking in such cases.

III. Can the co-citizen relationship be impaired?

In order to justify punishment, it is not enough that the victim and her co-citizens have the *standing* to view their relationship with wrongdoers as impaired. What if co-citizenship is

the sort of relationship that *cannot* be impaired? If, for example, the only feature of the co-citizen relationship was that all co-citizens are subject to the same coercive institutions, then though citizens might have the standing to view their relationship with the wrongdoer as impaired, the relationship could not become impaired. This is because wrongdoing does not stop the wrongdoer being subject to the same coercive institutions as the rest of the citizens. If this was the case, blame-based expressivism would not be able to justify punishment because no citizenship right could be suspended since the relationship could not actually be impaired. I will now explore some features of the co-citizen relationship that can be impaired by wrongdoing: solidarity and equal membership.

Equal Membership

I assume that liberal democratic societies have value, though I am not going to explicitly argue for this claim in the thesis. Central to the idea of liberal democratic citizenship is that all citizens are equal members of the community. This is a key feature of the co-citizen relationship and grounds many of our citizenship rights.

People might instinctively think that our status as equal members of our political community is grounded on the fact that we all have equal moral status. If this is true, then one might think this poses a problem for blame-based expressivism: our relationships with others can be damaged but (let's grant) our status as beings of equal moral worth cannot be. Does this mean that the state cannot deny our citizenship rights without denying our status as moral equals? I do not think so. My argument is that, whilst our status as moral

equals creates a presumption in favour of the idea that we should have equal membership, this does not mean that denying a person equal membership rights has to mean that his equal moral status is denied. Before I consider the particular question of whether we need to have equal membership in our political groups to have our status as moral equals respected, I shall make two points to motivate the general idea that equal memberships of a group is not needed to have our status as moral equals respected.

The first point to make is that our experience tells us that we do not normally think that we need to treat everyone as equal members of a group in order to respect them as moral equals. We do not normally think that non-members of our groups need to be given equal consideration in decisions made by a group to have their status as moral equals respected. For example, the British government does not need to give British and French citizens equal consideration in the decisions it makes in order to respect French citizens as moral equals. We also do not normally think that people need to be granted the status of equal member in every group they are part of in order to have their status as a moral equal respected. For example, an employer does not need to make all her employees partners in order to respect their status as moral equals (compare Miller, 2009: 295; Miller, 1997: 226-227; Griffin, 2001: 316-318).

The second point to make is more structural: the claim that we all have equal moral value is made from some objective perspective, but claims about the status of one's membership in a group are not like this. They are more based on the idea that there is a special connection between the parties in the group which means that there is value in those relationships being constituted by equality. This means that we can fail to grant someone

equal membership without denying their equal moral status. To illustrate let me take an example from a slightly different context: if I start a new relationship with someone, I will start to give them more consideration and special privileges than others, but this does not mean that I judge them to be of more objective value than other people or disrespect others' equal moral status (c.f. Scanlon, 2008). Likewise, people do not need to be granted equal membership of all groups on pain of having their objective value to be judged as less than it is.

So we can see that in general we do not need to be given equal membership of all groups to have our status as moral equals respected, but what about our political groups? Can we remove citizenship rights from wrongdoers without undermining their status as moral equals? I believe that we can. This is because blame-based expressivists claim that wrongdoers damage their co-citizen relationship. This means that the wrongdoer has to some extent removed himself from the group of citizens and therefore no longer a full (or equal) member of the group. Therefore just as when we do not deny the equal moral status members of other countries when we deny them rights of equal membership, we do not deny wrongdoer's equal status when we deny them rights of equal membership. (Wrongdoers are not fully non-members, like citizens of other countries, so the wrongdoers will have more rights in their country than them).

One might pose the following objection: in general it seems unacceptable for states not to give all their citizens equal membership. For example, it is unacceptable for Saudi Arabia to treat women as second-class citizens. The most obvious way of explaining this is to say that women and men are moral equals and treating women as second-class disrespects this

status. If this is the answer, then surely claiming wrongdoers do not have full citizenship status not denies them their equal status. Equal membership, the objector says, is not something that can be impaired because it is required by our equal status.

In response, nothing I have said thus far precludes me from believing that our equal moral status is relevant to why we should form democracies. In fact, I think our status as moral equals creates a strong presumption in favour of us having equal membership of our political communities. However, saying this is not the same as saying that we cannot deny wrongdoers equal consideration without disrespecting their status as moral equals. The fact that a person has equal moral worth in comparison to others is of course relevant to how we should treat her, and it is relevant to whether she should be granted equal membership in our political community. But there are many more variables to feed into our overall or general relationship with a person here: our history, the way we treat each other, how well we know each other, whether we share a language, etc. Therefore many variables feed into whether a person should be granted equal membership of our political community. What matters in determining whether we can non-objectionably suspend relations of equality with a citizen is whether she has damaged her relationship with co-citizens. She does not do this by virtue of her being a woman, but she does do this by committing an act of wrongdoing against another member. Therefore, on my view, it is objectionable for Saudi Arabia to treat women as second-class citizens because their equal moral status creates a presumption in favour of the idea that they should be part of a relationship of equality. The fact that they are women does not dislodge this presumption. Thus treating these women as second-class citizens does fail to take their moral equality seriously. However, it does not follow that we cannot suspend relationships of equality with any citizens without their status as moral equals being disrespected (new immigrants and ex-patriates do not possess

the same rights as everyone else and this does not seem to disrespect their status⁴⁵).

Therefore, it does not follow that we cannot suspend relations of equality with *wrongdoers* without disrespecting their status as moral equals.

Solidarity

Solidarity between citizens is another feature of the co-citizen relationship. Above I argued that citizens of a democracy should share we-intentions. A ‘we-intention’, recall, is an intention to do a particular activity *together* with another person or group of people. As I argued earlier we-intentions amongst citizens in a democracy involves viewing the democratic process as one where citizens gives each other’s interests moral weight rather than as a *modus vivendi*. Therefore citizens have solidarity with each in the sense that they are in it together when it comes to state actions.

If I commit a criminal wrong against my co-citizen it seems fairly obvious that I am undermining my solidarity with her. If I stabbed a number of my co-citizens, I do not act in a way that presupposed that we are ‘in it together’.

⁴⁵ Although it does when they are not accorded enough to meet their basic rights.

IV. Summary

In this chapter, I have sought to explain how we can apply the idea of relationship impairment to the relationship between citizens. I have argued that it is plausible to think that there are features of the co-citizen relationship which can be impaired by wrongdoing. These features are that citizens are equal members of the political community and that they share solidarity. Thus, the co-citizen relationship is the sort of relationship that can be impaired by wrongdoing. I have also argued that citizens' solidarity with each other and their position as each other's representatives give citizens the standing to blame on each other's behalf, they may all view their relationship with wrongdoer's as impaired. Thus, the idea of relationship impairment can be used to justify the claim that the state, rather than just the victims, may passively punish wrongdoers.

V. Implications and objections

Wrong Done to Non-Citizens

One might wonder about the implications of my theory for whether and how the state may punish wrongs done to non-citizens by citizens and, conversely, whether and how the state may punish wrongs done to citizens by non-citizens. I shall consider the latter question in the next section and the former question now.

If a group of citizens were to beat up a tourist or a non-citizen resident, would the state legitimately be able to punish them according to blame-based expressivism? I have argued that punishment is justified partly because features of the co-citizen relationship licence citizens to blame on each other's behalf. Citizens of state X are not co-citizens with tourists or non-citizen residents, so presumably the features of the co-citizen relationship that licence citizens of state X to blame on behalf of their fellow citizens are absent when it comes to tourists or non-citizen residents. So presumably blame-based expressivism implies that the state is not permitted to blame on their behalf and thus may not punish those who wrong them. If it were true that blame-based expressivism could not explain why the state may punish citizens who criminally wrong non-citizens, it would be a problem for the theory. I shall consider the case of non-citizens residents first, and then tourists.

When it comes to non-citizens residents, I think it is quite plausible to think that the state does have the standing to blame on their behalf and that this is perfectly consistent with blame-based expressivist ideas. It seems reasonable to assume that citizens have some level of solidarity with long-term residents. They may not have it with them to the same extent that they have it with their co-citizens, but, insofar as long term residents work in the community, are the sort of person considered eligible to receive benefits in the community, contribute to the state's political culture and obey its laws in a significant and ongoing way, and haven't committed any acts of wrongdoing against citizens, there is a real sense in which citizens are in it together with them. Similarly, though citizens do not represent long term residents in the same way that they represent each other, they do represent them to some extent: we expect the state's decision making body to pay special heed to their interests, more heed than they should pay to citizens of other states who do

not live in our state. This includes their interest in not being criminally wronged. This, I think, makes the wrong done to the non-citizen into the legitimate business of the state and can permit it to blame of their behalf. I believe that, taken together, these two points are enough to give citizens licence to view it as relevant to their lives when a long term resident is wronged and thus that they have the standing to blame on their behalf.

What now of tourists? Citizens do not appear to share any special solidarity or representative relationship with tourists, so can we justify passive punishing citizens who wrong tourists? I believe that we can, though the explanation has the disadvantage of suffering from the wrong-wrong problem. As I said above, I think that all citizens owe it to each other to obey the law. Therefore when a citizen breaks the law he wrongs all his other co-citizens. This gives all citizens some standing to blame him for that act and thus they have licence to punish. Again as I said above, this is not the most satisfying way of explaining the standing to punish because it suffers from the wrong wrong and wrong victim problems. It makes punishment solely a response to the wrong done to all citizens, rather than also a response to the wrong done to the victim. In cases where the wrong done attaches to a clearly identifiable tourist victim, it would be better if we could make our punishment a response to the particular wrong that the particular victim suffered.

However, I think in the case of tourists, we simply lack the relevant relationship to have the standing to do this. Perhaps this is as we should expect: if punishment is inherently a social practice (as I suggested it is in chapter 1), then perhaps we should not be surprised if the form it takes varies according to our social relationship with the victim. We should also remember that punishing citizens for wrongs done to tourists is only likely to form a fairly marginal set of cases. Thus, although I think it would be better if blame-based expressivism could avoid the wrong wrong and wrong victim problems when it came to

tourists, I do not think this deficit is a decisive consideration against the theory: we can still explain why the state has the standing to punish; perhaps we should expect things to be different when the social relationships involved are different; and we're only talking about a fairly marginal set of cases.

*Wrong Done by Non-Citizens*⁴⁶

As a representative of its citizens, the state has the *standing* to blame wrong done to the victim by anyone, non-citizens included. However, is there really any *space* for the state to manifest this blame? I have argued that blame-based expressivism permits the state to suspend citizenship rights, but if the state is not in a special relationship with non-citizens, it does not owe them any special citizenship rights. Thus it has no special duties to suspend in order to manifest its blame of non-citizens. So does the state have any way of meaningfully punishing non-citizens?

I believe that it does. I have claimed that citizenship rights include the right to move freely around a territory, and some political rights such as the right to vote. It is true that we do not give tourists voting rights and equal social rights and so we cannot remove these from them. However, we do usually allow them to move freely around our territory. As I will argue in chapter 5, they are not entitled to this so this privilege can be taken away to express our blame of them.

⁴⁶ I thank John Gardner for pushing me to think about this issue.

Wrongs Done to Dissenting Citizens

I argued above that citizens' solidarity with each other and their position as each other's representatives give citizens the standing to blame on each other's behalf. I also claimed that the co-citizen relationship one where solidarity obtains because running the state is a shared activity and, since democracy is more than a *modus vivendi*, citizens are 'in it together'. But what about citizens who deny that we are 'in it together' in the sense that generates solidarity? Do they licence the state to blame on their behalf? If not, may the state punish on their behalf? If not, then is this an unattractive implication for blame-based expressivism?

I think we can deal with the problem of dissenting citizens. First, it seems to me that the dissenting citizen can transfer the standing to blame to her state if she wants it to blame the wrongdoer on her behalf. Second, although a dissenting citizen may disagree that she is 'in it together' with her fellow citizens, we may nonetheless feel that the state still ought to represent her and this might give it the licence it needs to view wrongs done to her as its legitimate business. These considerations can plausibly give the state standing to blame and therefore punish on her behalf.

It might be the case that the dissenting citizen does not want to transfer the standing to blame to the state. She may be a principled anarchist and be unable to bear the thought of the state blaming another on her behalf, for example. In this case, there may be reason to think that the state may not punish on her behalf. As I said in chapter 3, I believe that there

is something attractive in giving citizens the power to waive the punishment of those that wrong them. Thus I do not think that the implication that victims can waive their wrongdoer's punishment is necessarily a problem. But, I am not sure that this implication follows. As I also argued in chapter 3, I am unsure whether victims do have the power to waive their punishment. I said that if your child went on to forgive someone who wronged her, it might not follow that you are then required not to engage in blaming behaviours towards him. Thus it might even be the case that the state has the standing to blame even when the dissenting citizen does not want to transfer the standing to blame to the state. Just because the citizen does not want the wrong done to her to be the state's legitimate concern, it does not mean that it is not the state's business. If the state represents its citizen, this might be the case whether she likes it or not. Therefore the wrong might be the state's legitimate concern regardless of a citizen's dissent.

Conclusion

In the last chapter I argued that being in a special relationship with a person gives you special rights against her. If you wrong her, this impairs your relationship. This means that the source of the special rights that you held against her is weakened. Therefore, depending on the extent to which the relationship is impaired, your special rights against her have been temporarily or permanently lost or weakened and it is therefore permissible for her to suspend the duties that correlated to these special rights. In this chapter I sought to defend the idea that citizenship is the sort of special relationship where this applies. If I am right then the state may suspend the citizenship rights of wrongdoers when it passively punishes them, though it may not suspend their basic rights.

However, though I have claimed that citizenship right include the right to move freely around a territory the right to vote, I have not yet defended these claim or explored the content of basic rights. However, the plausibly of my thesis depends on the content of basic rights and whether I can defend my claims about citizenship rights. The next chapter turns to these issues.

CHAPTER 5

Which Rights May The State Suspend From Criminal Wrongdoers?

Throughout this thesis I have argued that when a citizen commits a blameworthy criminal wrongdoing, she undermines her relationship with the state and so loses some of the special rights that were grounded in that relationship. I called these rights ‘citizenship rights’. I also claimed that we have rights to what we need for a minimally decent life and argued that state may not suspend these rights because they are grounded on our humanity rather than the fact that we are citizens of a state. I called these rights ‘basic rights’. In this chapter I argue that the right to move freely around the public spaces of a state is a citizenship right and therefore can permissibly be lost by wrongdoers. One of the main defining features of a prison is that it restricts the freedom of movement of inmates, so if the right to move freely around a public space is a citizenship right then blame-based expressivism has the beginnings of an argument as to why we may restrict the movement of wrongdoers by the use of prisons. I also argue that the right to vote is a citizenship right. Therefore blame-based expressivism may be able to permit the state to withdraw voting rights from wrongdoing citizens.

However, blame-based expressivism can only permit the suspension of the right to move freely around the state’s public spaces and the right to vote if doing so does not also undermine basic rights. Prisons obviously have the potential to do this and it is not unreasonable to think that removing the right to vote might too. In this chapter I argue that

this need not be the case. The first section of the chapter gives some terms and qualifications. The second argues that the right to free movement around a state is a citizenship right. The third section sets out what I think our basic rights are. The fourth sketches what a prison that respected these rights must be like. The fifth argues that the right to vote is a citizenship right and that losing it does not also violate basic rights. The last section considers objections.

I. Terms and Qualifications

(i) Because the state may not suspend basic rights, I will sometimes refer to these as *protected rights*. Because the state may suspend citizenship rights, I will sometimes refer to them as *unprotected rights*. Rights can be protected either because they themselves are grounded on our humanity or because they are instrumentally necessary to protect one that it is. In this way then, some rights which might initially appear to be citizenship rights, can be protected. For example, the right that our state (acting on behalf of our co-citizens) try to protect us against grievous bodily harm (through e.g. its police force) is a right needed to protect our basic rights that is held primarily against our own state and thus might initially appear to be a citizenship right.

(ii) Though I focus much of my discussion on our basic rights, they are just a subset of our protected rights. We have other rights that are grounded on our humanity. Such rights might include things like the right not to be lied to, rights against discrimination, and a right to due process. These rights may also be grounded on other considerations such as

respect or fairness. According to blame-based expressivism, the state may not suspend such rights because, even though they are not basic rights, they are not grounded in the special status of citizenship and thus wrongdoing does not undermine their grounds. To have a complete picture of blame-based expressivism, one would need to draw up a comprehensive list of protected rights. However, I restrict my discussion to basic rights because it provides a focused way of giving meaningful content to a substantial range of rights that we can be confident should be protected. They also seem to be the most relevant because these seem to be the rights most likely to be necessarily threatened by imprisonment.

(iii) The idea that special rights arise amongst co-citizens is not universally embraced. However, although I rely on this claim, I will not defend it. I also rely on the claim we may or should live in a world with states, which is also not universally embraced and which I will also not give a positive argument for (but see Miller, 2007; Wellman, 2016). This is because whether special duties arise amongst co-citizens and why we should give state's normative jurisdiction over their territories are enormous topics that I cannot adequately discuss. Also, even many of those with strongly cosmopolitan leanings believe that there are *some* special duties amongst co-citizens and even some of those who think that states should lack significant jurisdiction over who enters their borders are not always committed to the idea that they lack jurisdiction *within* them (e.g. Carens, 2016). So it is more relevant to argue about whether and why particular rights, including the right to free movement within a territory, are citizenship or basic rights, rather than whether there *any* citizenship rights or whether states should have normative control over their territory.

(iv) My discussion is not about *human rights*. Although there is some overlap with the literature on human rights and some philosophers do define human rights in a similar way to how I have defined basic rights (e.g. Nickel, 2005; Miller, 2007, 2016), I wish to avoid this term. This is because the phrase ‘human rights’ is now often used in an extensive way to include rights that only arise by virtue of citizenship. For example, the phrase ‘human right’ is usually used to mean a right that all persons in the world should have. We might think that all persons have a right to be members of a particular political community. Therefore some rights one has in virtue of being a citizen *could* be included within this conception of human rights. For example, the right to vote is a citizenship right that one might think everyone should possess but only because we should all be part of one political community or another (compare Miller, 2007: 168; Cohen, 2006). Also human rights doctrine sometimes seems to be addressed at states, telling them what rights they should give their citizens. Clearly these might be citizenship rights. Therefore since the distinction between rights that one possess in virtue of citizenship and those that one does not is important in my thesis, it is better to avoid the term ‘human rights’.

(v) Our *basic rights* protect our basic needs. Our *basic needs* are what we need to have fulfilled if we are to lead *minimally* decent lives.⁴⁷ It should be emphasised then that our basic rights are not that extensive - they do not give us what we need to live a flourishing life. It is important to include the ‘minimally’ in ‘minimally decent life’. This is because basic needs are meant to possess a certain moral urgency (Cohen, 2006: 229; Altman and Wellman, 2009: 3; Miller, 2007: 198). This is what makes them so widely accepted. It is also important that we try to constrain the extent to which basic needs stretch because if

⁴⁷ Please note that I am using the phrase ‘X has a basic need for P’ merely to indicate that X needs access to P if she is to have a minimally decent life. Thus I am departing from the way the term ‘basic needs’ is sometimes used in the literature (it is sometimes used to refer only to physical needs).

we start claiming that people need, say, full political equality or strongly egalitarian distributions of resources their societies to live minimally decent lives, then we are committed to the conclusion that say Jane Austen could not have had a minimally decent life. But, whilst her life may have been restricted in some very important ways, the idea that it could not have been minimally decent is implausible and, in fact, insulting (see Buchanan, 2015).

Drawing the line between what counts as minimally decent and not minimally decent without just asserting what one feels to be minimally decent or not is a difficult task (c.f. Buchanan, 2010: 691; Griffin, 2008: 3). We all need to assert things sometimes and I will have to do this to some extent in what follows. However I hope to minimise the extent to which I do this by basing my claims on a small number of underlying principles (that we have basic physical, autonomy-based, and social needs), which I hope have wide appeal. I will expand on these ideas and give my account of basic needs in section III, but first I will explain why I think that the right to move freely around a territory is a citizenship right rather than a basic right.

II. The beginnings of an argument that blame-based expressivism can permit prisons.

In this section I will give what I think are persuasive reasons to support my argument that the right to move freely around the public spaces of a state is a citizenship right, but I am aware that this is a controversial idea and there will be those who fundamentally disagree. These people will have to conclude that my claim that blame-based expressivism permits

prison is false. But, if they otherwise like my theory, they could explore some alternative version of blame-based expressivism so long as they think that some special citizenship rights exist. There are three main reasons why one might disagree with my claim that the right to move freely around the public spaces of a state is a citizenship right. These are that it is *itself* a basic right, that it is needed to protect some further basic right, or that it is some other kind of protected right. I think these ideas are mistaken and I will explain why. But first I will give a positive argument in favour the idea that the right to move freely around the public spaces of a state is a citizenship right.

If my friend and I bought a house together on equal terms, we would clearly be able to set up private areas (such as bedrooms). However I think it is reasonable to claim that, absent special agreements or considerations, we should both be able to freely use (or at least have equal access to) any areas that we have deemed as shared. Citizens of the state are all supposed to be equal members of the community that has jurisdiction over the territory. I think this more or less equates to the claim that all the citizens, in some loose sense, all equally own or as a collective have ultimate control rights over the territory. Though citizens may have private property, I think it is plausible to assume that their equal membership of the community that has jurisdiction or ownership of the territory gives them a claim to wander freely (or at least have equal access to) its public spaces. If a citizen were denied this, I think she could quite reasonably claim that she was not being treated as an equal member of the state.

If we think that a person has ownership or jurisdiction over some private property, we usually think that this means they have normative control over who enters that property.

We also usually think that they have, to some degree, normative control over what others may do or where they may go within that property. For example, if my friend and I have a guest over, we are perfectly entitled to request that he only enters the living room. Thus if we accept that citizens have normative control over their territory, we should also accept that they control of whether non-citizens may enter, and what they may do and where they may go if they are permitted to enter. Thus I believe that citizens have a right to roam freely around (or at least have equal access to) the public spaces of their territory but non-citizens do not. If I am correct about this, then in principle the state can suspend the right to move freely around the territory from wrongdoing citizens, because they have undermined their relationship with other citizens and thus no longer have the status of equal members of the community. This then opens the way for permitting the state to restrict the movement of wrongdoing citizens through prisons because prisons characteristically restrict free movement around the public spaces of a state.

It is important to note that having ownership or jurisdiction over something does not entitle a person to complete normative control of it. I own my book, but this does not entitle me to whack you over the head with it. If my friend and I own our house, this fact does not entitle us to ambush our guest in the living room and tie him to a chair. In general, our normative control over what we own or have jurisdiction over is limited by the rights that others hold against us, especially their basic rights. Thus it would follow that citizens' normative control over their state's territory does not permit it to violate the basic or other protected rights of others. Therefore, if the state would violate a wrongdoer's basic or other protected rights by a particular method of controlling their free movement within the territory, this would not be permitted. (This, of course, is fully in line with the rationale of blame-based expressivism, which does not permit the state to withdraw basic or other

protected rights). But this does not undermine the general idea that the right to move freely around the public spaces of a state is a citizenship, rather than basic or other protected, right. In order to undermine this idea, one would have to show that restricting free movement inside the territory necessarily violates or is very likely to violate basic or other protected rights.

First I shall consider the claim that that restricting free movement necessarily violates or is very likely to violate basic or other protected because the right to free movement is *itself* a basic or other protected right. I do not think this is a very plausible idea. This is because of the type of value free movement has. Following James Griffin, I think that freedom in general, and the right to free movement in particular, has little intrinsic value: as he says, stopping a person from driving one way down a street, choosing their meal for them, or imposing dress codes for certain dinners are restrictions on freedom. But they do not really seem to deprive anyone of anything of great value (Griffin. 2001: 320-321). The main value of freedom is in what it lets us do. If there is not very much value in a free movement for its own sake, then it cannot itself be a basic or other protected right because these rights are supposed to be urgent or otherwise important in themselves.

If free movement is not itself a basic or other protected right, then perhaps it is a basic or other protected right in the sense that it is needed to guarantee more fundamental basic or other protected rights. I shall discuss this momentarily but first we should note that no one who accepts private property can argue that nearly unlimited free movement is needed to protect our basic or other protected rights. Insofar as we think that a significant area of the state may be taken up by private residences or business, we accept that fairly significant

restrictions on free movement do not violate people's rights (c.f. Miller, 2016: 24; Wellman, 2016: 88). Of course, all our basic and other rights are limited by the rights of others. But they are not conventionally understood to be the sorts of things that can be persistently and pervasively curtailed in the way that free movement is. Some current distributions of private property are unjust, but in principle I do not believe that the institution of private property is. Therefore I do not think that it is plausible to think that nearly unlimited free movement is needed to guard our basic or other protected right.

So the question becomes, 'how much free movement is needed to protect our basic or other protected rights?' Some philosophers believe that *extensive free movement* is required. That is, they think that free movement between and within territories is required. I shall now consider whether such extensive free movement is needed to protect our more fundamental *basic* rights (I shall comment on other protected rights later). Kieran Oberman argues that people have an 'essential interest' in extensive free movement to protect both what he calls 'attachments' and 'possibilities'.⁴⁸ Attachments are 'the options that we have chosen or have somehow become attached to: our family, our career, our religion, and so on' (Oberman, 2016: 41). For reasons I will elaborate on in later sections of the chapter, I think we can have basic rights to access some *non-substitutable* attachments, that is, attachments that we cannot easily switch for others. Some access to one's religion and loved ones is necessary for a minimally decent life and we cannot just swap these things for other equally loveable people or another equally well-followed religion. Thus if someone is in a place where she cannot access her loved ones or religion we cannot justify a restriction on her freedom of movement simply by saying that she can access a different

⁴⁸ For another argument that extensive free movement is a human right see Philip Cole (Cole and Wellman, 2015: chpts 12 and 13)

religion or person to love where she is. This might plausibly undermine her basic rights. However, cases like this are fairly marginal, and do not mean that in general we all need a basic right to extensive free movement.

What of attachments that can easily be substituted for others? There are many attachments that we can discard altogether or substitute for others and still have a minimally decent life. For example, if one is attached to rock climbing, but finds oneself in a location where one must swap one's extreme sports hobby to kayaking, diving, or surfing it seems implausible to insist this would stop your life being minimally decent. So long as a person has access to an adequate range of substitutable options within a locality, it seems implausible to claim that restricting her movement to this locality would violate her basic rights – surely her life would still be minimally decent.

Possibilities are 'all the other options that we are not attached to but could, at least potentially, be interested in pursuing, either now or in the future' (Oberman, 2016: 41). If we had a basic right to this, then extensive free movement would be needed. But, although we clearly we need some possibilities to have minimally decent life, we most certainly do not need to be able to pursue all of them. Oberman claims that we must have the freedom 'to join and establish new civic associations, meet new people and make new friends, learn about and convert to new religions, and attend meetings on subjects we know little about' (Oberman, 2016: 44). He thinks that this is because, amongst other things, we have an essential interest in having the freedom to 'search for answers to questions of ultimate meaning' (Oberman, 2016: 44). But this is an implausibly luxurious conception of what

we should have basic rights to (c.f. Wellman, 2016: 89)⁴⁹ and even if it were not it seems implausible that we need extensive free movement to be able to learn about new things and pursue some new interests. Thus I do not believe we need to have extensive free movement to protect our basic rights.

Exactly how much free movement is needed to protect our basic rights is an open question, but it clearly hinges on what options are available within your locality. If a person's locality has an adequate range of options, then limiting her movement to this locality does not violate her basic rights. This is not to say that limiting peoples' movement to small localities is always permissible or never violates basic rights. Historically, people's basic rights have been violated by restricting their movement to certain localities. For example, the basic rights of those contained in Nazi controlled ghettos were clearly violated. The use of ghettos also violated residents other protected rights such as the right against discrimination as well as their citizenship rights. Thus blame-based expressivism is not committed to the implausible position that it is always permissible for state to restrict the movement of citizens or non-citizens.

What I need to know is whether states can restrict the movements of wrongdoers (who have undermined their status as citizens) to prison without undermining their ability to lead a minimally decent life and thus violating their basic rights. I believe that many current prisons do deny basic rights to inmates and thus do not grant them a minimally decent life.

⁴⁹ I should say that Oberman does use different terminology from me. However, I do not think I misrepresent him. Oberman argues that we have human rights to a *full* range of options, rather than a basic rights to a an *adequate* range of options (Oberman, 2016: 39-44). But he claims that human rights guard our 'essential interests' (Oberman, 2016: 39, my emphasis). Thus, we can make the same point: this is too luxurious an interpretation of 'essential'.

However, I believe that it is possible to have prisons that do not do this. They would need to plan the activities they offer inmates very carefully and offer opportunities for prisoners to see family and enter the community in a way that is uncommon for British prisons at the moment. In the next section I explain what I think our more fundamental basic rights are and in the section after I apply my ideas to suggest what a prison that respected basic rights would need to be like. However, what is important for now is that there is no principled reason to think that a prison must violate wrongdoers' basic rights.

Of course, it is open for the reader to disagree with my account of basic rights and argue that they must be much more extensive than I claim and, as such, could not possibly be upheld by an institution that we could plausibly call a prison. But it is worth noting the implications of making this claim. The idea that criminals retain their basic rights is widely accepted. If prisons necessarily violate basic rights, then we either have to reject the claim that criminals do not lose basic rights⁵⁰ or we have to radically reconceptualise how we think of state punishment.

Now let us consider the claim that extensive free movement is needed to guard a different kind of protected right, rather than a basic right. Probably the most popular and plausible protected right this might be is a right against unfairness (see e.g. Carens, 1987). Where one is born is a matter of brute luck and has a huge impact on one's life chances. Many people believe that it is unfair for something which is a matter of brute luck in this way to have such an impact on one's life chances and thus think that people have a right to

⁵⁰ I defend the general idea of rights forfeiture. But this does not commit me to the idea that wrongdoers forfeit basic rights – one of the problems with the strict reciprocity view that I reject is precisely that it seems to specify that wrongdoers can forfeit basic rights, for example, rights against rape and torture.

extensive movement between areas to guard against such unfairness. But accepting this claim does not show that the state may not punish wrongdoers through restricting their movement. This is because it is not a matter of brute luck that someone has blameworthily committed a crime and therefore, although restricting their movement through prison would impact their life chances, it would not be unfair in this sense.

III. Basic Needs

In this section I discuss what our basic needs are and some of the rights they ground. We cannot justify the loss of these rights by appeal to blame-based expressivism and therefore any prisons that were permitted under this theory would have to uphold these rights. In the next section I will sketch out what such a prison might be like. This sections needs to be quite long because we can't reach many conclusions about prisons and the form that they can take without having delimited quite carefully the nature and extent of the needs that ground basic rights.

I believe that people have basic physical, autonomy-based, and social needs. I follow other philosophers in believing that these needs must be met in order for people to have a minimally decent life. All seem to agree that basic physical needs, such as the need to avoid pain and hunger, make the list (e.g. Griffin, 2001; Fabre, 2000; Miller, 2007 Southwood and Barry 2011). Some philosophers, for example James Griffin and Cécile Fabre, also endorse the idea that that we must have our autonomy respected to some degree in order to live a minimally decent life (see Griffin, 2001; Fabre, 2000). Examples

of philosophers who also believe that we need some level of decent human contact and social inclusion if we are to lead a minimally decent life are Kimberley Brownlee, David Miller, and Fabre (Brownlee, 2013; Fabre, 2000; Miller, 2007). I now discuss each of basic physical, autonomy-based, and social needs in more detail.⁵¹

Physical Needs

Human beings have basic physical needs which must be met if we are to have a minimally decent life. At the most extreme level, if our most fundamental biological and physical needs are not met, we will die. But the physical and biological needs that must be fulfilled if we are to have a minimally decent life extend beyond what we need to survive: we need to be free from physical suffering. One's ability to live a minimally decent life can be undermined by, for example, persistent hunger, pain, cold, heat, and disease even if one is not, strictly speaking, at death's door (c.f. Fabre, 2000: 21; Miller, 2007: 180-181). When we see images of starving people or the terminally ill, for example, it is not just the fact that their lives are in danger that make us feel they are not enjoying a minimally decent life,⁵² it is our knowledge that they are in persistent hunger, pain or severe discomfort (c.f. Fabre, 2000: 21). One does not need to be totally free from pain, hunger etc to lead a minimally decent life and there will be some disagreement over exactly what degree of pain, hunger, disease etc. one needs to be free from if one is to enjoy a minimally decent life. But this does not undermine the general point: pain, hunger, thirst, and lack of

⁵¹ Though I divide basic needs into physical, social, and autonomy-based, needs note that there is some overlap between them.

⁵² Saying that these people are not living a minimally decent life is not the same as saying their lives are not worth living. I am sure that many of them would dispute the latter claim, but not the former, since the former only amounts to saying that persons have a right to be free from the suffering they are experiencing simply in virtue of their personhood (if this is achievable).

protection from the elements can stop one from having a minimally decent life (c.f. Miller, 2007: 182).

Social Needs

In the context of discussing basic needs, Miller and Brownlee have observed that, as well as being physical and biological creatures, people are fundamentally social and that we have basic social needs⁵³ corresponding to this fact. Miller says, ‘human beings are social as well as biological creatures, and they can be harmed by being denied the conditions of social existence’ (Miller, 2007: 182). Brownlee says that persons have a right against ‘a persisting lack of minimal opportunities for non-threatening, decent, or supportive social interaction’ (Brownlee, 2013: 206). I believe that Brownlee and Miller are correct on this score and I argue that we have basic social needs for some decent social contact, to not be forcibly prevented from involvement with our loved ones, and for some level of what I call social functionality. I shall now examine each of these ideas in more detail.

a. Decent social contact

Brownlee’s primary concern is with the need for bare social contact, that is, our need to actually interact with others. Brownlee narrows down the contours of the need for social

⁵³ I should be clear here, that the social needs I am discussing are not the same as the ‘societal needs’ Miller identifies. My social needs are simply needs for social contact and inclusion. Miller’s societal needs are basic needs that only arise in particular social contexts and, as such, they vary from society to society.

contact. It is not a need for just any old social contact: ‘hostile, degrading, or cruel’ contact will not do (Brownlee, 2013: 200). Nor is it a need for friendship or love (Brownlee, 2013: 206). Nor is it simply a need not to be interfered with in our attempts to make contact with others. Instead, it is a need to be provided with ‘minimally adequate opportunities for decent or supportive social contact’ (Brownlee, 2013: 207).

It is undeniable that we need opportunities for decent social interaction in order to live a minimally decent life. I shall summarise two of Brownlee’s arguments for this conclusion, both of which I find convincing. First is her ‘normative argument’: she believes that good social interaction is intrinsically valuable, ‘having adequate social contact in the form of supportive interpersonal interactions, acceptance, and associative richness is...valuable in itself’, so valuable that it is a necessary constituent of a minimally decent life (Brownlee, 2013: 203, 212). In defence of her claim, Brownlee points out that individuals who choose to spend extensive periods of time in isolation - such as astronauts, solo explorers, nuns, and monks - often experience the loneliness of their chosen path as its most challenging part, and it often takes considerable training for them to be able to endure it (Brownlee, 2013: 205). Moreover, we might note that these individuals probably do have some opportunities for fairly regular social contact. This is certainly true of nuns and monks and, as Brownlee mentions, astronauts often become dependent on regular radio contact (Brownlee, 2013: 205).

Second, is Brownlee’s ‘empirical argument’. A review of the psychological literature on loneliness finds that it ‘generates the same threat response as pain, thirst, hunger, or fear by setting off a chain of anxiety-inducing physiological reactions known as the ‘fight or

flight’ response’; and is associated with depression, increased systolic blood pressure, ‘the progression of Alzheimer’s disease, increased vascular resistance, diminished immunity, reduced independent living, alcoholism, suicidal ideation and behaviour, over-expression of genes bearing pro-inflammatory responses, and mortality in older adults’ (Brownlee, 2013: 211). Adults subject to solitary confinement ‘report feeling depressed, lethargic, forgetful, and then mindless, sleeping over twelve hours a day and lacking the energy to read, eat, or move. Some people hallucinate, have panic attacks, mutilate themselves, or become semi-catatonic’ (Brownlee, 2013: 212). Many people who have experienced long-term solitary confinement testify that they found it as disturbing and agonising as any physical abuse they suffered (Brownlee, 2013: 212). Thus it seems clear that some level of decent social contact is needed for a minimally decent life and thus there is a basic right to it.

One might object to this claim, however, for similar reasons to why you might object to the claim that people can have a right to love or friendship. We cannot have rights to love or friendship because, as Brownlee herself argues, these relationships ‘cannot be produced on demand’, so even if there were a basic need for these goods it would be irrelevant. This is because it would not be able to ground a right since there cannot be a corresponding duty (Brownlee, 2013: 206). But the right to decent social contact (from a state official) is not the same as the right to friendship or love. It only requires *supportive* contact, which can be produced on demand. It is similar to the sort of contact we might think a vulnerable child is entitled to receive from a social worker or teacher: respectful treatment designed to support and acknowledge them. For wrongdoers this could be provided by prison wardens or councilors as a duty of their job, not as an act of friendship or affection.

b. Freedom to maintain relationships with our loved ones

As I said, we do not have a basic right to be provided with friendship or love. But— where these relationships already exist - I think we do have a basic need not to be forcibly prevented from maintaining them, especially when it comes to our closest relationships such as those between parent and child and domestic partners.

Here is some evidence for this claim. Mothers in prison report ‘feeling anger, anxiety, sadness, depression, shame, guilt, decreased self-esteem and a sense of loss when separated from their children’ (Byrne and Stanley, 2000: 3). Women’s anxiety over the care of their children can damage women’s mental health (Revolving Doors Agency, 2002). Women have also been found to feel depressed because they miss the support provided by their families (O’Keffe et al, 2010). In male prisoners, being unable to keep up family ties has been linked to mental instability and it is widely recognised that prison visits are essential to maintaining a prisoner’s mental well-being (Lindquist, 2000).

Putting aside these empirical considerations surrounding forcible separation from one’s family, it seems reasonable to assume that maintaining (good) personal relationships is of intrinsic value: many people would cite it as being the most important thing in their lives. Thus, it seems to me that being forcibly removed from one’s closest relationships undermines a person’s ability to lead a minimally decent life. Evidence for this can be found in the fact that, (i) in the ordinary run of things, it seems that it would be totally unjustifiable to forcibly separate someone from her family and (ii) this would be true even if the person was not a co-citizen (most states accept, and we think they should accept, that

the ‘right to a family life’ should be interpreted as requiring that people should be permitted residency in their spouse’s country).

I believe that we have a basic need not to be forcibly prevented from seeing our loved ones. We do not need to be able to fully participate in our family life to have a minimally decent life, but we certainly need to have enough contact to keep the relationship going and to prevent mental instability and illness. It would be implausible to understand this as a need to see our loved ones on a near-daily basis – many people take periods away from their families to, for example, further career goals. But people think it is desperately important that they are aware of and involved in the development of their young children, which may well require near-daily phone or video calls. Domestic partnership typically revolves around very high levels of intimacy where each party is aware of the day-to-day details of the other’s life, but domestic partners can probably maintain their relationship and minimally decent lives whilst being forcibly prevented from sharing this level of detail. Nonetheless, to maintain relationships of this type we need to be allowed multiple opportunities for phone or video calls in a week. Both types of relationship, especially parent-child, also require some opportunities for face-to-face contact. For people who lack these kinds of relationships, the support networks provided by other family relationships or friendships may be just as highly valued and important for maintaining mental stability and thus should be protected in a similar way.

Of course some people do elect to leave valuable family relationships and some people do not have them (because their family treat them badly or have died or left etc.). But we should note that being forcibly prevented from seeing or having regular contact with one’s

loved ones is quite different from electing to leave. Most relevantly for our discussion (because this is often true of prisoners), this seems especially plausible when one is also put in an environment where one cannot compensate for the loss by pursuing some other valued goal or relationship.

c. Minimal social functionality

I believe our social needs extend beyond the need for social contact and the need to maintain our special relationships. There is a broader sense in which we need to be protected from social exclusion: as social creatures, I believe we have a need to have a place in the public life of our society and to be able to play some part in it. I shall capture the idea that we have a basic need to be able to participate in society by saying that we have a basic need for ‘minimal social functionality’.

I think it is best to understand the need for social functionality as a need for some level of recognition from the other members of your society. It is a need for recognition of the fact that you are a fellow human being and not some object that is simply present in the society. The need for basic social functionality is not a need to be able to participate in society as an *equal*, since one can have recognition as a member of a group and a human without having recognition as an equal member of a group (see below). But one does need to be able to engage to some extent in the public culture of one’s society, and, in this sense have a role in it: one cannot be totally excluded from the public life of one’s political

community. I think the opportunity for this sort of recognition is intrinsically valuable and necessary for a minimally decent human life.

As far as I am aware, the need for minimal social functionality has not explicitly identified before. But, other authors do hint at it and thus I think it is an idea that would receive support. For example, James W. Nickel believes that, in order to live a minimally decent life, people need to be able to participate in ‘the *social*, political, and economic’ activities of their community (Nickel, 2005: 388). Fabre talks about the need to be able to ‘function’ in one’s society and also states that one of the conditions for having adequate access to a range of decent options is ‘not being excluded through lack of means from the social fabric and from the public space in which [the options] present themselves’ (Fabre, 2000: 12). Some of Miller’s examples of basic needs seem to make reference to the idea that we have a basic need to be able to play a part in our society. He also believes that feeling ‘socially excluded’ is something that can cause your life to be less than minimally decent (Miller, 2007: 182).

Therefore I believe that people have a basic right to minimal social functionality: they have a right to engage in a minimal way in the public culture of their society. This means that they need *some* opportunities to enter public spaces and participate in other ways in the public culture (perhaps through writing in public discussion forums). To have a minimally decent life, a person does not need to be able to access the public culture whenever they like or on equal terms to everyone else. But they need to feel some level of recognition from their culture and not to feel totally excluded.

Autonomy-Based Needs

I shall define autonomous choices as those where a person is pursuing something she, in some sense, thinks is valuable, worthwhile or, as it is commonly put, forms part of her conception of the good. Autonomy is often understood to mean not merely that a person is acting out their conception of the good, but that they've arrived at that conception through rational reflection, or through aligning it with higher-order desires. I should signal now that the way in which I use the word 'autonomy' is more down-to-earth than this. It seems unnecessary to me that a person has consciously and rationally worked out a conception of the good and that only the actions that tie directly to that should count as autonomous. Most of us are not that reflective. I think autonomous actions are those which the author genuinely values, but a person need not have consciously and rationally worked out a full conception of the good to be able to pursue actions that she genuinely values.

I believe that we need to have some autonomy if we are to lead a minimally decent life. We do not need to live a *fully* autonomous life to lead a minimally decent life – we do not need to be able to succeed in pursuing the best life we think is possible, nor do we need to spend all our time pursuing our conception of the good. So the questions that need to be considered in order to provide an account of how much autonomy we have a basic need for are 'what type of actions count as pursuing your conception of the good?' and 'how much do we need to be able to pursue these actions in order to have a minimally decent life?'

I will consider the ‘what type’ question first. Both Fabre and Griffin believe that protecting autonomy is important because it protects something that we think is valuable and sets us apart from other animals – people’s ability to reflectively decide how to pursue their life (Fabre, 2000: 18; Griffin, 2008: 157). Despite not agreeing with the extent to which the decisions must be reflected on, I agree with this general sentiment. Therefore the sorts of actions that count as pursuing your conception of the good need to be something more than just doing something that feels nice (like eating chocolate ice cream) or that are just a basic response to a stimulus (like shifting in your chair). The sorts of actions that count for most people will be things like: pursuing personal relationships, hobbies that one values such as drama or painting or film, career, and religion.

Many of the actions that we perform, however, do not count as pursuing our conception of the good. These include some of the actions orientated towards our survival (Fabre, 2000: 10-11), maintaining our standard of living, and meeting social norms. Most of these actions are instrumentally necessary to pursue our conception of the good, and sometimes these types of actions can be constitutive of our conceptions of the good. But often we do not need to count them as examples of actually pursuing it. For example, though actions orientated towards my survival, such as eating and drinking, are instrumentally necessary for me to pursue my conception of the good, it would seem implausible to say that I am pursuing my conception of the good when I drink a glass of water. Conforming to social norms is also necessary, or at least makes it a lot easier, for us to pursue our conception of the good. But we need not count all examples of maintaining social norms as pursuing our conception of the good. For example, suppose I value my career, but I do not value having to stick to the norms of personal attire that form part of my company’s ethos. It does not seem right to think that I am pursuing my conception of the good if I nonetheless

unwillingly stick to it. It may sometimes be difficult to draw an exact line between actions that are constitutive of our good and those that are not. (For example, are the late nights and intensive rehearsals necessary to put on an exciting piece of theatre something which adds richness and value to the creative process or are they a necessary evil to be endured?) But this does not undermine the point that many of the actions that are orientated towards our survival, maintaining our lifestyle, or meeting social norms are not themselves examples of pursuing our conception of the good.

There are also a myriad of trivial decisions that we make everyday – decisions about when to eat, what dress to wear, whether to get up and make a coffee. Often performing these actions does not need to count as pursuing your conception of the good. We can see this by considering the fact that having some constraints on our trivial choices does not on the whole undermine our ability to form and pursue a conception of the good. As Griffin says, if your mother insists on choosing your meal when you go to a restaurant, or your college insists that you wear a gown to dinner, it hardly stops you from pursuing your conception of the good (Griffin, 2001: 320-321). Clearly there must be limits on the constraints on our trivial choices. If all, or nearly all, of our trivial choices were controlled then this *would* start to undermine our ability to pursue a conception of the good. This is because many trivial choices might be thought to express our identities and our identities constitute part of our conception of the good. Therefore, extremely persistent or pervasive constraints can undermine our autonomy based needs, but this does not change the fact that many of the trivial actions that we pursue need not be thought of as pursuing our conception of the good.

Now I am going to consider *how much* we need to be able to pursue our conception of the good to have a minimally decent life. Fabre claims that we need to have a decent *range* of options and adequate *access* to these options (Fabre, 2000: 10-13). I agree with this and wish to go into a little more detail as to what this means. I will consider what it means to have decent access first. One thing that Fabre argues is necessary for decent access is ‘not being forced to take an option which does not leave us time to do anything else’ (Fabre, 2000: 11). As an example of to illustrate this point, let me use an example adapted from Fabre and Raz: A woman is trapped on a desert island. There are plenty of things to do on this island (hiking, botany, exploring etc.) however the woman lacks any meaningful autonomy because food is so scarce that she must spend all her time either hunting, gathering, eating or sleeping. ‘Everything she does is geared towards preserving her life, and every single choice that she makes is literally a matter of life and death’ (Fabre, 2000: 11). This woman has no time to pursue her conception of the good and therefore has virtually no meaningful autonomy and clearly is not living a minimally decent life.

This tells us that we need some time to pursue our conception of the good to live a minimally decent life. But how much? Many of us spend a lot of our time performing actions that are trivial or orientated towards survival or meeting basic social norms: washing, dressing, cleaning our homes and our clothes, brushing and drying our hair, tidying, traveling to work, boring administrative jobs that form part of our work, (for some people even) working itself, eating, food shopping, cooking, waiting for things etc. It is true that some of these activities will form a constitutive part of the good life for some people. It is also true, for reasons I gave above, that others having too much control over these actions might start to undermine our ability to lead a minimally decent life. But the general point remains – most of us spend a lot of time doing things that we don’t really

consider as a constitutive part of our good life. We may do this and yet consider our life to be rich and full of many valuable activities. In short, we can spend a lot of time *not* pursuing our good life and yet live a much more than minimally decent life. So it is my contention that perhaps we only need to spend a few hours a day pursuing our conception of the good to have just a *minimally* decent life. For example, imagine that the woman on the island only had to spend her morning and part of her evening searching for food and was at liberty to spend the afternoons pursuing other interests such as hiking and botany. Assuming that the woman's basic needs are met and there are no further bars to her accessing decent options in the afternoon, it is plausible to assume that she has enough autonomy to lead a minimally decent life.

What now about the idea of needing a *range* of decent options to be autonomous. We touched before on the types of actions that might count as examples of pursuing your conception of the good. I said we might include things like pursuing personal relationships, hobbies that one values such as drama or painting or film, career, and religion. Fabre says 'human beings have dispositions to take up jobs and hobbies requiring different skills and preferences: manual deftness, intellectual abilities, artistic abilities, a taste for sports, outdoor activities, and so on. A decent range of opportunities should include opportunities of these kinds' (Fabre, 2000: 11). It is fairly obvious if a person had no choice of valuable ends to pursue; if there was only one that was available to her she is very likely⁵⁴ to lack meaningful autonomy. But we do not need to be able to do whatever we have decided constitutes a good life, or pursue it in whatever way we like, in order to have meaningful autonomy. It is worth discussing substitutable and non-substitutable

⁵⁴ I say very likely because if it just so happens that the type of life open to her was the one she would have chosen anyway then perhaps she is autonomous.

activities separately. Let's discuss substitutable activities first. We tend to value a plurality of these activities, so we can still have meaningful autonomy if we cannot pursue all the substitutable goods we value. For example, if I value basketball, football, and going to the pub, I can still pursue a life of meaningful autonomy if I only have the opportunity to pursue two of these goods in my leisure time. Therefore we can still pursue a life that we think is valuable from a constrained set of substitutable options so long as some of the things within the option set fall into what we consider to constitute a good life. Of course we could not constrain this set too heavily – perhaps we would need one or two options of different types, for example, sport, culture and social. This would not be a fully autonomous life, but it could be a life with enough autonomy to have a minimally decent life.

That said though, this is not the same when it comes to non-substitutable attachments, for example pursuing special relationships and religion. So whatever range of options we offer, we must include access to these. However, we do not need to allow people to pursue non-substitutable goods in whatever way they want in order for them to have a minimally decent life. For example, perhaps Hannah views following Christianity as a necessary to the good life. Maybe she wants to attend the church that she grew up attending and would like to attend Wednesday and Sunday mass. But she can't do this. Instead she has to attend a different church, which only offers Wednesday mass and she can still read the bible and try to live according to Christian values as far as possible. She is not living her ideal good life, but her ability to pursue this element of her conception of the good is not undermined to the extent that it stops her from living a minimally decent life. There are limits to this –the provision of the category of the good that she thinks is essential must meet some standard of adequacy. Hannah's church must, say, not be overrun by

corruption so that she is prevented from attending an organization that genuinely pursues Christianity.

So, in this section I have considered what types of actions count as pursuing your good and the extent to which they need to be protected in order to preserve the autonomy needed for a minimally decent life. I have claimed that trivial and instrumentally valued actions do not always need to be considered as examples of pursuing your good life. This means that we do not need to spend more than a couple of hours a day pursuing our conception of the good in order to have a minimally decent life (because all of us spend a lot of time not pursuing our goods whilst having more than minimally decent lives). We can achieve a minimally decent life when we only have a constrained set of substitutable options and though we need access to our non-substitutable attachments, we do not need to be able to access them however we like.

IV. Prison under blame-based expressivism

In section I, I argued that the right to move freely around the public spaces of a state is a citizenship right and therefore, according to blame-based expressivism can be suspended during punishment. I said that this gave us the beginnings of an argument that blame-based expressivism could justify prisons because the main characteristic feature of imprisonment is that it confines you and restricts your freedom of movement. But blame-based expressivism can only permit prisons if, by restricting free movement, they do not indirectly violate any other protected rights. Is this possible?

Before answering this question, however, it is worth commenting on the enormous variety there is between different prisons and different individuals' experiences within prisons. English prisons are divided into four categories (A-D), with A being for the most dangerous prisoners who are deemed most likely to try to escape. A category A prisoner may not be allowed to wear his own clothes or leave his cell unless there are a specific number of prison officers on the wing, whereas a category D prisoner may have his own clothes, have keys to his own cell and be on day release. Within prison wings of the same category inmates have different privileges depending on their behaviour. For example, an inmate on basic privileges may be allowed only the legal minimum of visits and not be allowed a TV in his cell, whereas a prisoner on enhanced privileges may have earned extra visiting privileges, time out of cell, a TV and a Playstation. Different prisons have different prison rules, offer different opportunities to inmates, have their cells open for different lengths of time, and are managed with different levels of success. This variation between different prisons and prison experiences shows that when we are considering whether blame-based expressivism can permit prison, there is not one fixed institution that we must have in mind. There is scope for permitting institutions of quite different kinds which might all plausibly be considered to be prisons.

I shall now sketch out what a prison that allowed prisoners to have a minimally decent life might look like by suggesting what it could do meet each of the basic rights discussed above in turn.

Basic Physical Needs

In order to be permitted by blame-based expressivism, a prison would clearly have to meet the basic physical needs of inmates by providing them with food, water, shelter, and basic medical care etc. This is not controversial and is in line with current practices. Under the European Convention on Human Rights prisoners have a right to food, water and healthcare and most UK prisons aim to deliver these rights.

So, with respect to rights to have basic physical rights met, blame-based expressivism goes with the *status quo*. There is, however, perhaps one exception here: though the death penalty is outlawed in all Western democracies except the USA, it does enjoy a considerable amount of popular, and some philosophical, support (Kramer, 2011). A blame-based expressivist system of punishment would firmly deny that the death penalty is an appropriate form of punishment because it clearly violates basic rights.

Social Needs (Decent Social Contact)

I claimed above that people have a right to a certain basic level of decent or supportive social contact. Many prisoners do have access to this. They have access to prison chaplains, phones, some will form positive relationships with specific prison staff, most prisons offer 'listener schemes' etc. However, solitary confinement is used in many Western democracies (though few governments use this term now). The US makes prolific

use of the practice: tens of thousands of US inmates are held in solitary confinement at any one time and, indeed, it is the norm in US supermax facilities (with most states possessing at least one prison with supermax capabilities). In Britain a standard ‘in prison’ punishment is cellular confinement for up to 14 days (Loucks, 2000).

Our basic need for social contact does not ground a right to unlimited social contact. But prisoners under solitary confinement in the US are typically confined to their cell for between twenty-two and twenty-four hours per day, which is clearly unacceptable. In the UK, a prisoner governor, a member of the chaplaincy, and a healthcare official will visit each prisoner in solitary confinement daily. Prisoners in solitary confinement may also be given some association time with other inmates, but this is not a legal right. Whether this is sufficient to ward off the psychological damage caused by isolation is an empirical question that I cannot answer. But, it still leaves the solitary prisoner in isolation for long periods of time. It seems plausible that this might render his life non-minimally decent and we should be very wary of believing that a person’s need for social contact can be met under these conditions. Therefore this type of imprisonment would not usually be permitted under blame-based expressivism, at least not for any protracted period of time.⁵⁵

⁵⁵ Solitary confinement is often claimed to be a necessary form of protective custody, either to protect the segregated prisoner or to protect others. In line with what I have said elsewhere, I believe that if solitary confinement really is necessary for protection, it may be used. But I believe these cases are extremely rare, every effort should be made to ensure some decent contact is maintained, and we should compensate the wrongdoer for this because it is a (permissible) rights infringement.

Social Needs (Freedom to Maintain Relationships with our Loved Ones)

I claimed above that people have a basic right to be able to maintain a meaningful level of involvement in their special relationships with loved ones. I said that this required frequent access to phone or video calls as well as some opportunities for face-to-face contact.

Prisoners' need to maintain their relationships with their loved ones is widely acknowledged in English prisons. In terms of frequent phone calls, the prison service instructions⁵⁶ state that inmates should be given access to a telephone at times when friends and family are available and not normally for less than two hours per day. In practice it is difficult for prisoners to spend two hours a day on the phone to loved ones because of demand for prison phones, availability of loved ones, and costs. But the kind of phone contact required in a prison permitted by blame-based expressivism would not be markedly different from how it is supposed to be now.

In terms of face-to-face contact, the average UK prisoner is permitted three to four sixty-minute visits per month (though they are only legally entitled to two). The nature of prison visits and visitor areas varies from prison to prison. In some prisons they take place under sterile and heavily monitored conditions, that allow very little physical contact and hardly permit prisoners to spend so-called 'quality-time' with their families. The short and sterile nature of these visits can impede a prisoner's ability to maintain their relationships with loved ones: prison blogs are full of warnings about how disappointing and awkward visits from loved ones can be. However, this is not universally the case. Some visitors' halls are pleasantly decorated and have children's play areas and visitors' cafes; some prisons allow

⁵⁶ These are the regulations and guidelines according to which prisons should be run.

hugging and hand holding, which allows for a more relaxed and ‘normal’ interaction.

There is a move towards allowing ‘special visits’ in prisons where inmates can meet with their families in a more informal setting for longer periods of time and participate in some activities together such as sharing a meal, watching a film, or playing sport. HM Prison Holloway (a women’s prison) used to offer family day visits to enable mothers to spend a decent length of time with their children (Rickford, 2003:10).⁵⁷ Conjugal visits are not currently permitted in the UK but many states do allow them.⁵⁸ Canada allows all inmates (except those deemed at risk of family violence) ‘private family visits’ of up to 72 hours every two months. These can be attended by anyone who has a close familial bond with the inmate. It is also worth noting that many English category D prisoners are permitted to go on home visits, which must make it much easier to maintain special relationships as it would allow much more relaxed and intimate interactions.

I suggest that in order to make sure that an inmate’s ability to live a minimally decent life is not undermined by being prevented from maintaining their special relationships, prisons must maintain the kind of phone access they are meant to already have and must allow inmates to spend regular (perhaps weekly) quality time with their special relations by allowing them to meet for longer and in the more informal settings mentioned above. A prison that allowed this kind of interaction between inmates and their special relations might be more liberal than some people would be happy with. But it would still recognisably be a prison and it still has great expressive significance to put someone there because of a blameworthy wrong they have committed. Therefore, at least with regards to

⁵⁷ Sadly these were cancelled due to staffing issues.

⁵⁸ For example: France, Spain, Saudi Arabia(!), Mexico, Brazil, Denmark, and India as well as some US states and some Australian territories.

the right to maintain special relationships and basic physical rights, it seems that prison may be permitted by blame-based expressivism.

Social Needs (Minimal Social Functionality)

In section III, I said that I believe that people have a basic right to minimal social functionality: they have a right to engage in a minimal way in the public culture of their society and to some level of social recognition. Thus they should not be totally excluded from public life. Since a defining feature of a prison is that it confines inmates, this right will be the most challenging for a prison to accommodate. However, I believe it is possible. In order to still recognisably be a prison, I think an institution would probably require the residency of inmates. But that does not mean that they cannot be let out or that opportunities to access public culture cannot be brought in.

A prison could help inmates to fulfil their need for social functionality in different ways. One would be to arrange social trips and activities for prisoners in the outside world, the other would be to allow more people from the outside world to come and interact with prisoners. Although trips out are not the norm at higher security prison categories, most prisoners progress through the categories and end up in category D prisons. Prisoners are then eligible for Release on Temporary License (ROTL) to work, go on home visits, or visit the local town. They become eligible for ROTL after passing their Full License Eligibility Dates, which is usually only a quarter of the way through their sentence. Therefore the number of prisoners who are currently able to leave UK prisons may be

higher than one might expect. If we want to ensure that prisoners obtain minimal social functionality, it is probably better to allow prisoners some opportunities to go out. However, opportunities can be brought in. An example of the latter kind of endeavour was 'Barbed', a graphic design business established by the Howard League for Penal Reform inside Coldingley Prison. The business had to be closed due to legal technicalities but it clearly gave workers some opportunity to have a role in the economic and public life of the UK. Another current example is the UK's National Prison Radio, based in Brixton Prison, where prisoners work alongside outside professionals to deliver award-winning programmes to the prison population. The Clink is a charity that runs high calibre restaurants in four UK prisons. These restaurants are open to the public. Inmates are trained by professionals and interact with the public in the restaurants. This again would give inmates the opportunity to have some minimal social functionality. We could also give them simpler ways to interact, such as, blogging, letter writing, or limited time windows in which they can access social media (some of which are already allowed).

I believe that prisoners need to have opportunities such as the ones mentioned in the above paragraph to have their right to minimal social functionality respected. I tentatively suggest that prisoners should be allowed one opportunity a month to either leave the prison or work with an outside organisation within prison. Therefore blame-based expressivist prisons would need to afford prisoners these opportunities. This might, to some extent, require a shift in some people's everyday understanding of what 'prison' means. But I think an institution that afforded prisoners opportunities to enter the community or have the community come to them could still be recognisable as a prison so long as it required their residency, that they spent most of their time there, and that it still put many

constraints on what they may do (which it would, see below). After all there are examples of such institutions (category D prisons) already within our justice system.

One obvious concern over letting prisoners out would be that some prisoners are too dangerous to be let out into the public. Another concern would be that prisoners might escape. To mitigate concerns about escape we could specify that some outside visits would need to be supervised by prison staff. However, a more cost effective solution could be to monitor prisoners by using electronic tags and satellite surveillance (as is already done for those under Home Detention Curfew). In terms of public safety, we might point out that reoffending rates amongst those on Home Detention Curfew are dramatically lower than for those released without tags, suggesting that tags might prevent those released into the community for day release from committing crimes during this time; that reoffending rates are lower for those on high level community service orders than they are for those serving custodial sentences; and that it is widely recognised that maintaining strong links with community and family helps reduce recidivism. All of this leans in favour of the idea that releasing prisoners into the community may actually be better in terms of long term safety.

However, there may be some prisoners who really are so dangerous that it would be justified to restrict their entrance into the community. But, we must note that this cannot play a large role in justifying locking up our prisoners. First, if we do subject these prisoners to full-lock up, this would have to be a permissible rights infringement. That is, wrongdoers have not forfeited their right against it so it would still wrong the wrongdoer by infringing their rights (even though it is all-things-considered permissible). We do not want to have a situation where we are relying heavily on the idea of permissible rights

infringements to justify imprisonment: permissible rights infringements are supposed to be rare occurrences. Second, such dangerous criminals only make up a very small part of the prison community. It is difficult to find information about exactly how many they number, but here is some information to get an idea. There are 130 prisons in the UK. 9 of these are classed as category A prisons. This just means that they have *some* capacity to hold category A prisoners and often only a small proportion of the prison's population is in this category. For example, Belmarsh prison has a total capacity of 855, but in 2014 only 51 (6%) of prisoners were categorised as category A and, within that class, only 11 were deemed 'high risk' and only 3 'exceptionally high risk'. Therefore, even though there are some prisoners who it may be best not to allow out, they may well not be very numerous.

Autonomy Based Needs

In section III, I said that a wrongdoer should be able to pursue her conception of the good to a certain extent. She must be able to pursue some activities that she thinks have value, and she must have some access to her non-substitutable attachments, but she does not need unconstrained options or unconstrained access to those options. I said, however, that these constraints are subject to the condition that the options are offered to an adequate standard (and not, say, corrupt or badly managed). I also said that we can control another's trivial choices, but that there must be *limits* on the constraints on our trivial choices. If all or nearly our trivial choices were controlled then this would start to undermine our ability to pursue a conception of the good by, for example, not allowing us to express our identities. I also said that many people who live much better than minimally decent lives spend a lot of their day engaged in activities that do not plausibly count as pursuing their conception of

the good. Therefore we do not need total freedom in our daily routine to have a minimally decent life.

Can a prison allow inmates to pursue their conception of the good to a sufficient extent to allow them to lead minimally decent lives? There is huge variety in the normal regime (timetable) in different UK prisons but the general pattern is as follows: inmates eat breakfast in cells out of a weekly issued breakfast pack; inmates have a morning period of exercise, work, and/or education;⁵⁹ inmates return to cells with ‘controlled unlocked time’ where they collect lunch from the canteen and eat in cells; inmates spend an hour or so locked in cells; inmates have an afternoon period of exercise, work, and/or education; inmates return to cells with controlled unlocked time to collect dinner from the canteen and eat in cells; inmates have one or two hours of association time (where prisoners can shower, call home, etc).⁶⁰ The government target is that prisoners should spend 10 hours a day unlocked in the week. Conditions are usually more restricted at weekends. Prisoners can attend religious services on the appropriate day for their religion.⁶¹

When they are in their cells, prisoners have the right to enough things to fit into two boxes, plus a larger music playing device (radio, guitar etc). They can also have -though do not

⁶⁰ <http://insidetime.org/regimes/>

⁶¹ In practice, funding cuts, staff shortages, violence, and overcrowding mean that prisons often fail to uphold their normal regimes. There are not enough work and education places and not enough staff to supervise unemployed prisoners to be out of their cells during the morning and afternoon periods. Exercise periods too are often cancelled due to lack of staffing and bad weather so that in practice an unemployed prisoner may well spend above 20 hours in his cell each day. However, we want to see if blame-based expressivism *could* justify something that would be recognisable as a prison so I leave this to one side and will suppose that a normal regime is running.

have right to - a TV, Playstation, extra reading and writing materials, and their own bedding. Many trivial choices will be taken away from prisoners. However, they typically do retain some. For example, most prisoners are allowed to wear their own clothes, they have coffee and tea rations and kettles in their rooms, have a choice of radio and TV stations, and they can buy items such as chocolate from the prison canteen.

So do prisons allow prisoners to meet their basic autonomy-based needs? When it comes to substitutable goods, much of this depends on what options they offer during the work, education, or sports times. Many prisons try to offer a wide range of activities. For example, HMYOI Rochester lists Art, Basic Education, Computer Studies, Cookery, Crafts, Creative Writing, English, Key Skills, Languages, Life and Social Skills, Literacy, Maths, Music, Numeracy, and Open University as available courses. HMP Belmarsh works in partnership with Charlton Athletic FC to offer coaching programmes and lists the following as 'available sports': Badminton, Basketball, Circuit Training, Light Circuit Training, Over 40s, Over 50s, Remedial, Soft Tennis, Sports Field, Volleyball, Weight Loss Programme, and Weight Training. When it comes to non-substitutable attachments, I have already discussed access to loved ones. Another main area for concern is access to religion. All prisons allow inmates to pursue their religion, some even order in special foods for religious festivals.

If a prison is able to ensure that the provision of these options meets a basic standard of adequacy, then, so long as these activities are carefully planned, it seems reasonable to assume that inmates would be able to pursue some non-substitutable options that they value. I think that this sounds like a prison, can in principle, allow prisoners to meet their

basic autonomy-based needs, especially when we consider that they would also have to allow the access to community and loved ones that I detailed above. In practice, many prisons may not be able to ensure that provision of the opportunities *does* meet a basic standard of adequacy is a big one and effective prison management would be needed to ensure it, but this doesn't undermine the point that one could in theory have one's autonomy based needs fulfilled in prison. Prisons permitted by blame-based expressivism must provide these opportunities to their inmates, and these institutions would be clearly recognisable as a prison.

Throughout my thesis I have argued that the state may suspend citizenship rights from wrongdoers but that it may not suspend their basic rights. Thus far in this chapter my main aim has been to consider whether blame based expressivism can permit prisons. I argued that freedom to move around the public spaces of a state is a citizenship right, not a basic right and that therefore the state may be able to permissibly imprison wrongdoers if it can do so without also violating their basic rights. I believe it can do this by ensuring that wrongdoers' basic physical, social and autonomy-based needs are met. It can do this by, for example, not using solitary confinement, allowing longer and more informal family visits, carefully planning prisoner activities, and allowing prisoners to leave the prisons for electronically monitored visits.

I'm sure a system that accords the opportunities I've discussed to inmates would strike some as very lenient and radical. However, I think that those who would want to defend prison terms that are harsher than the ones I've been describing would need to do so on desert-based retributive grounds but I rejected these in chapter 1. I also hope the

information I have provided about the opportunities that are (or theoretically should be) already available to prisoners in the UK and other states indicates that the type of prison permitted by blame-based expressivism is not actually as radical as it might first appear.

IV. The Right to Vote

The right to vote is indisputably a citizenship right in the sense that it is a right held only against one's co-citizens (we do not expect the USA to allow UK citizens to vote in its elections) and it only arises because we stand in a special relationship of co-citizenship with others. Thus according to blame-based expressivism, the right to vote is unprotected. However some may object to this idea. The first form of objection is to say that though the right to vote is a citizenship right in the senses just mentioned, it is also a protected right because, although it does arise out of the context of citizenship (or other group membership), this context is an enabling condition of the right rather than its grounds. Its true ground, we might think, is this fact that all persons are beings with equal moral status – this is the real reason we think people should have the vote. The fact that they are part of a political group just provides the relevant context for the right to vote to come into play. If this is the case, then perhaps the right to vote is protected. This would be because wrongdoing does not undermine our status as beings of equal moral value so, according to this line of objection, wrongdoing does not undermine the grounds of our right to vote. However I do not think that this line of objection is successful. I do think that our status as moral equals creates a presumption in favour of the idea that we should be given political equality, including the vote, in the political groups that we are equal members of. But, as I argued in chapter 4, I do not think it means that we must be granted equal membership of

all groups. For example, it is not an affront to the moral equality of persons if French citizens are not given equal membership in the UK. When blame-based expressivists deny wrongdoers the vote, they are not denying their status as persons of equal moral worth, they are claiming that the wrongdoer no longer has the status an *equal member* of the political group. This is because the claim that blame-based expressivists are making is that the wrongdoer has damaged his co-citizen relationship. This means that the wrongdoer has to some extent removed himself from the group of citizens. This is not the same as denying his status as a being of equal moral worth. Thus I believe that the idea that it is our status as beings of equal moral worth that grounds our right to vote, does not mean that the right to vote must be protected.

It does not follow from this that denying people equal membership to a political group is never an affront to their status as beings of equal moral worth. Many people have been denied equal membership of their political communities in a way which is an affront to their status as moral equals. Let's assume that British women were denied the status of equal members of their political community because they were thought to be irrational and/or infantile, or were thought to lack interests separate from their fathers/husbands. These are examples of failing to respect women's status as moral equals. Denying women equal membership was thus an affront to their status as beings of equal moral worth and objectionable. But it does not follow that denying wrongdoers equal participation must affront their status as people of equal moral worth. We can quite consistently say that wrongdoers can lose rights to equal political participation whilst still being moral equals.

The second line of objection against the idea that the right to vote is an unprotected right is that it is needed protect our basic needs. However, I do not think this objection holds. I shall now explain why it is (i) not constitutive of autonomy based basic needs, (ii) not constitutive of the need for social functionality and (iii) not required for more instrumental reasons.

Autonomy-Based Needs.

It is tempting to think that democratic systems are autonomy-enhancing for individuals in the sense that voting can be thought to be an extension of a person's decision making from the individual realm to the political and that the ability to do this might form part of someone's conception of the good. This, however, is deeply misleading. As has been capably argued by others before me, the fact that a law or a person has been democratically elected does not make our laws the result of each individual's reflective decisions about questions of the good. As Griffin says, in a modern democracy, one's vote makes no difference to the policies selected and so does not enable one to live in the way one has reflectively decided upon: 'to have one vote among millions does practically nothing to protect one's being able to pursue one's chosen ends' (Griffin, 2008: 248). Indeed the policy selected might hamper one's attempts to pursue one's conception of the good. Wellman and Altman also and similarly illustrate that democracy is not autonomy enhancing in the sense that it extends individuals' decision making. They emphasise that it is the group, rather than the individual, that decides the law in a democracy. The fact that the group will decide the answer to a particular question takes an individual's power to decide the answer to that question herself out of her hands: 'democratic decision-making

on a given matter is inconsistent with each individual's possessing autonomy on that same matter' (Wellman and Altman, 2009: 21).

Nonetheless there is *a* sense in which democracy can be autonomy enhancing. For many people, the decision to vote for one party rather than another is the outcome of reflective deliberation about which values should be placed front and centre in public life and which party is most committed to these values. It can also be an important expression of identity. However, we do not need the opportunity to exercise our autonomy in this way for our autonomy-based *basic* needs to be met. We do not need to be able to exercise our autonomy in this way to ensure that we have the opportunity to pursue some activities that we value and to have access to our non-substitutable goods.

The need for Social Functionality

The right to vote is a right to be recognised as an equal by one's co-citizens. But recall that one does not need to be able to participate in a society as an equal in order to have one's need for basic social functionality met: one just needs to have the opportunity to play a part large enough to mean that one has some recognition and one has not been excluded from social life. One can get achieve minimal social functionality by participating in the kind of activities I discussed in section III. However, we might reasonably think that our political life forms part of our social life, so we might reasonably think that the right to social functionality encompasses a political, as well as other social aspects. I agree with this, but one can still achieve the right to social functionality without the right to vote. One

can get recognition in the political realm through other activities such as writing in political discussion forums and drafting and signing petitions, for example. It might at this point be useful to distinguish between a right to vote and a right to political participation and claim that we have a basic right to political participation, but not a basic right to vote.

More Instrumental Reasons

One might argue that the right to vote should be protected for instrumental reasons: the best way of ensure that citizens' basic needs are met is to insist that they should have voting rights. Griffin, for example, rejects the idea that the right to vote is strictly necessary for a minimally decent life but he believes that, in the real world, we do need it to protect basic needs:

[It is not] empirically possible for a society in typical modern conditions of highly pervasive government, large population, advanced technology, concentration of coercive power at the centre, ethnic diversity, educated citizenry, with the sort of social cohesion necessary for a tolerably successful democracy, and so on, to respect all [basic needs], yet not be democratic (Griffin, 2008: 249) .

Wellman and Altman appear to endorse a similar view:

We readily admit that (in the real world, at least) the absence of democracy tends to render [the fulfilment of basic needs] considerably less secure (Altman and Wellman, 2009: 34)

We could interpret these claims in two ways. The first is that democracy in general plays a role in protecting basic rights. On this interpretation particular groups might benefit from

being in a democracy even if they themselves do not have the right to vote. I agree with first interpretation but clearly it does not mean that wrongdoers must be given the vote as they can benefit from being a democratic society without being given it. The second interpretation is that particular groups need the vote if they themselves are to have their basic rights protected. I agree that the right to vote has been instrumentally useful in protecting the basic rights of particular groups in society, for example women. However, it is less clear whether wrongdoers make up one of these groups. Prisoners can vote in the Czech Republic, Denmark, Finland, Ireland, Latvia, Lithuania, Macedonia, Montenegro, Serbia, Spain, Sweden, Switzerland and Ukraine but not in the UK, Bulgaria, Estonia, Georgia, Hungary, Liechtenstein, Russia and, until recently, Austria.⁶² It is not clear whether prisoners in the latter group of countries are significantly more likely not to have their basic needs met than those in the former, and, even if it were, it is not clear that we could attribute the difference to voting rights. Thus it seems that the first interpretation of the argument is the more reliable one to base our judgements on.

V. Objections

Can the Right to Free Movement really be Lost?

Above I said that perhaps the reason citizens have the right to move freely around the public spaces of her state is that all the citizens, in some loose sense, own or have collective control rights over a the territory One might object that if this is the case then a

⁶² Data correct November 2012.

citizen does not lose her ownership of the public spaces when she commits a crime because this is not how ownership rights work. For example, if you and I buy a house together and then I undermine our solidarity by beating you up, I do not lose my ownership rights over the house. The way I lose my ownership rights over the house is if I sell them or give them away. So why would a citizen lose her right to move freely around the public spaces?⁶³

In response: the way in which one gets ownership rights affects the way in which one can lose them. In the above example, you get your ownership rights by buying the house. You do not get them in virtue of your relationship with me, so of course your damaging our relationship does not make you lose your ownership rights over the house. However things are not always like this. If I am given free membership of a tennis club I have claims to walk around it and use its facilities. But if I get kicked out for bad behaviour these claims are lost. Consider now citizens' claims of ownership over the state's public spaces. These are more similar to my claims at the tennis club than your claims over our house. Citizen's claims have not been bought. They are based on membership of the state – non-citizens do not have these claims. Therefore it seems to me that these claims can be lost if you damage your citizenship status. Of course, things are a bit more complex in the case of the citizen than in the case of the tennis club member. Perhaps some of the citizen's claims are also grounded on factors like reciprocity or birth-right and so perhaps the citizen does not fully lose her claims to use the public spaces. However, assuming that solidarity also plays a part, and then this just provides another argument in favour of

⁶³ We might interpret this as a version of the objection that blame-based expressivism confuses fittingness with entitlement that I discussed in chapter 3.

letting wrongdoers out on visits into the community - does not mean that the force of her citizenship rights cannot be weakened.

Does Blame-Based Expressivism Imply that the State could Permissibly put Non-Wrongdoing Non-Citizens in Prison?

As we have seen, blame-based expressivism says that wrongdoing citizens damage their relationship with the other citizens and thus that other citizens may suspend their rights of citizenship, such as the right to free movement and the right to vote. But what about non-citizens such as long-term residents, those on short-term working contacts, and tourists? If they do not possess citizenship rights, then does this mean that blame-based expressivism implies that it would be *prima facie* permissible to imprison them even if they have committed no crime? That is, would blame-based expressivism imply that imprisoning innocent non-citizens does not violate their rights? If so, would this be an unattractive implication? I will consider the case of long term residents first, then short term residents and tourists.

When it comes to long-term residents my response is that, as I said in chapter 4, I think it is fairly plausible to assume that they gain similar rights to citizens. It seems reasonable to assume that citizens have some level of solidarity with long-term residents. They work in the community over long periods of time, are the sort of person considered eligible to receive benefits in the community, contribute to the state's political culture and obey its

laws in a significant and ongoing way.⁶⁴ Thus it seems plausible to me that long-term residents do to all intents and purposes become citizens or very close to citizens and should acquire citizenship rights or a set of rights that are very close to citizenship rights. Thus it seems reasonable that long term residents possess a quasi-citizenship right to freely roam the public areas of the state. Therefore blame-based expressivism would not imply that it is *prima facie* permissible to imprison them.

However, what about people who come here on short-term working contracts or on holiday? It seems less likely that they would gain something close to citizenship rights. Thus, blame-based expressivism does suggest that it would be *prima facie* permissible for the state to imprison them – that it would not violate their rights to put them in prison. This may at first seem to be a troubling implication but I think it can be mitigated. The main reason for this is that blame-based expressivism does not imply that it would be all-things-considered permissible for the state to put innocent non-citizens in prisons. This is because there would be no positive reason to put them there since there is no expressive value in imprisoning innocent people. There are also weighty reasons against doing so. Examples of reasons against imprisoning them might be: it would cost a great deal to imprison and monitor them; their lives would go better if they were not in prison and generally speaking, when there is no cost to us or other reason against allowing someone to lead a better life, we should not stop them; imprisoning short-term workers would likely discourage people from coming and bringing economic benefits; it would prevent them from enriching the cultural lives of our community; and we have granted them permission to come and live or work in our state and this may plausibly mean we have some fiduciary duty to grant them

⁶⁴ For an argument that long-term residents have fairly similar (although not identical) rights to citizens based on a mixture of reciprocity, solidarity, and coercion see Sarah Song (2016).

temporary more extensive rights than simple basic rights.. Another reason that the implication that it is *prima facie* permissible for the state to imprison non-citizens is that anyone who recognises the territorial authority of states is in any case committed to a similar conclusion: tourists and temporary residents are here at the whim of the state and thus do not possess rights against detention and removal (at least assuming that detention and removal does not compromise their basic rights).

Conclusion

In this chapter I argued that the right to move freely around the public spaces of a state is a citizenship right and therefore can be lost by wrongdoers. One of the main defining features of a prison is that it restricts the freedom of movement of inmates. Therefore blame-based expressivism can restrict the movement of wrongdoers by the use of prisons if it can do so without also violating their basic rights. I believe it can do this by ensuring that wrongdoers' basic physical, social and autonomy-based needs are met. I think that ensuring that prison meet these basic rights makes for a liberal conception of a prison, but nonetheless something which is still recognisable as one. I also argued that the right to vote is a citizenship right that can be suspended from wrongdoers and this can be used as a form of punishment. Some of the arguments I have given have been controversial, but the arguments of the chapter mark a step forward for rights forfeiture and expressive views because they explain how two forms of treatment that we typically associate with punishment can be justified. They also mark a step forward for theorists of punishment in general because these theorists rarely consider what forms of punishment their theories justify and whether these forms are acceptable.

I have now finished giving the positive arguments of my thesis. The next chapter is devoted to defending blame-blame based expressivism from a series of objections that could be raised against theories of punishment that make a person's blameworthiness relevant to why the state may or should punish them.

CHAPTER 6

Problems with Thinking of Punishment as a Response to

Blameworthiness

In chapter 3, I distinguished between the judgment that a person is blameworthy and blaming a person. I said the judgement that a person is blameworthy for an action is attributing moral fault to her for that action and evaluating her in light of it. I said that blaming a person involves judging that she is blameworthy and viewing your relationship with her as impaired. But I haven't said anything about what sort of fault a person needs to possess in order to be blameworthy. Once we start to look into what it takes to be blameworthy, we might worry about using a blame-based account of punishment. I shall begin by saying a little about what it means to be blameworthy, before explaining why using blameworthiness as a basis for punishment might lead to problems.

The judgement that a person is morally blameworthy for an action presupposes that she is agent-responsible for it. A person is agent-responsible for action V if we can attribute it to her in the sense that she is the author of it: that it has arisen in some appropriate way from her agency because she, in a relevant sense, endorses it. Deciding whether something has arisen from a person's agency in the appropriate way is very complex, but it might be useful to contrast it with causal responsibility. With causal responsibility, we only need establish that an agent or object has a relevant causal link with action V to say she or it is causally responsible for it. So for example, if Mary smashes a glass when sleep walking, we may say she is causally responsible for it even though she is not, on most accounts,

agent responsible for it (because she is too radically isolated from agency to attribute authorship of smashing the glass to her).⁶⁵

We must be agent-responsible for a wrongful action to be blameworthy for it: Mary is not blameworthy for smashing the glass. But the judgement that a person is agent-responsible for an action is not enough to show that she is blameworthy for it. This could be because a person can be agent responsible for a morally neutral action, like intentionally standing on one leg, or praiseworthy actions like being moved to help a stranger for the right kind of reasons. One might think that if a person is agent-responsible for a morally wrong action, then she is blameworthy for it. However, this is not always the case. We know this because an action's wrongfulness can come apart from our judgements of blameworthiness. For example, we can sensibly say that a person's action was wrong but that we shouldn't blame her because her heart was in the right place or because she was reasonably mistaken about something (c.f. Scanlon, 2008: 124-5).

When we judge that a person is morally blameworthy for an action, part⁶⁶ of what we are doing is judging that the way she has authored it shows that her moral deliberations about why to pursue that action were *morally faulty*. That is, when we say a person is blameworthy we are judging that her reason for action is itself an immoral reason and/or that she has been negligent by failing to give the correct moral weight to a pertinent moral

⁶⁵ I'm indebted to Stemplowska and Knight (2011: 11-12) for structure and example inspiration in this paragraph.

⁶⁶ I say '*part* of what we are doing' because it may also be necessary for us to judge that a person's action is wrong or bad in a more objective sense in order for us to judge that she is blameworthy. For example, if a wealthy person begrudgingly gives to charity in order to make her company look better and therefore garner more business, it is somewhat strained to say she is blameworthy even though her deliberations are somewhat morally faulty. But I leave this open.

reason for or against acting (c.f. Arpaly, 2006: 14). For instance, imagine that attending my child's 6th birthday would make her very happy. If I fail to attend because I reasonably but falsely believe my mother has been rushed to hospital, I am not blameworthy. My reasons for action are not immoral nor have I given a reason the incorrect moral weight in my deliberations. If I fail to attend in order to try and make my child unhappy, I am blameworthy because I do the action for the very reason that it is wrong: the reason I do it is an immoral reason. If I fail to attend because I would rather watch Coronation Street, I am also blameworthy. Wanting to watch Coronation Street is not in itself objectionable, but prioritising Coronation Street over making my child happy in this case indicates that I haven't given my moral reasons for action the appropriate kind of weight in my deliberations.

We can now begin to appreciate the sort of concerns one might have about viewing state punishment as a response to blameworthiness. First, if we start to separate a person's blameworthiness from the wrongness of her action, do we end up in a situation where the outcome of a person's action becomes irrelevant to how much we punish her? For example, imagine that Tom shoots Tilly and Tim shoots Tabitha. If both men act as a result of equally faulty moral deliberations, should we punish them equally if Tom kills and Tim misses? In other words, can blame-based expressivism allow for the phenomenon (of something very similar to what) Nagel dubbed resultant⁶⁷ moral luck (Nagel, 1979)?

⁶⁷ Resultant moral luck concerns how luck affects the outcomes of actions. There are other types of moral luck such as luck in the character traits one possesses (constitutive moral luck) and the circumstances in which one finds oneself. I do not consider these kinds of moral luck here.

Second, we can be blameworthy for all sorts of actions. This includes actions that we would typically think of as private or as trivial. It may even include actions that pose no risk of harm and even mere thoughts. If blame-based expressivism is a response to a citizen's blameworthiness, must a proponent of it be committed to punishing these types of actions (and thoughts)?

Third, if judging that a person is blameworthy involves judging that her moral deliberations about why to take an action were morally faulty, then blame-based expressivism would, obviously, require the state to judge and punish a person on the basis of her moral reasoning. One might think that the state has no business assessing a person's deliberations because, for example, it is an invasion of privacy or a restriction on freedom of thought.

In this chapter, I respond to these worries. In section I, I show that blame-based expressivism actually deals fairly well with moral-luck cases (indeed one of the touted attractions of Scanlonian blame is that it has the resources to do this). In section II I deal with trivial and private wrongs and non-harmful actions and thoughts. In section III I deal with the objection relating to the state assessing a citizen's moral deliberations.

I .Accommodating Moral Luck

Moral luck, as I define it, occurs when one agent can correctly be subjected to more moral criticism than another agent despite the fact that they have both displayed an equal level of faulty moral reasoning.⁶⁸ Consider the following examples:

Molly: Molly has a reckless disposition (that is, she does not care sufficiently about the impact of her actions on others). One day she is driving particularly recklessly. Because of her recklessness, the car comes very close to hitting a child called Sarah. But through sheer good luck she does not injure Sarah.

Matthew: Matthew has the same reckless disposition as Molly. One day he is driving with the same level of recklessness as Molly is in the above example. He hits a child called Stephen as he crosses the road. Were it not for his recklessness, he would not have killed Stephen.⁶⁹

If moral luck exists, then Matthew can correctly be subject to more moral criticism than Molly. If it does not, then he cannot. Our current practices indicate that most of us accept moral luck. However, Matthew and Molly have equal faulty deliberations so it *seems* that accounts of blameworthiness, such as my own, which tie a person's blameworthiness to their deliberations and tie relationship impairment to blameworthiness, must be committed to denying that moral luck exists. If this were the case, then a proponent of blame-based expressivism would have to be committed to the idea that individuals like Matthew and

⁶⁸ Nagel defines it as follows: 'Where a significant aspect of what someone does depends on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck' (Nagel 1979, 59). But, because I believe that we can meaningfully hold a person blameworthy even if determinism is true, I wish to try avoid the concept of control as much as possible (here I follow e.g., Arpaly, 2006; Scanlon 2008: chpt 4)).

⁶⁹ Nagel/Williams (1976) made these kinds of examples famous. I model the versions in the text more closely on Scanlon's (Scanlon, 2008:148-149)

Molly should receive equal punishment. At the level of intuition, I do not find it obvious that this would be a decisive consideration against it. After all, there is something clearly unfair about treating Matthew and Molly differently. Despite my intuitive sympathy with denying that moral luck exists, however, I do believe that a compelling theory of blame must be compatible with moral luck. In part, this is because rejecting it is so radical – divorcing the actual outcomes of our actions from how we are treated for them would be bewildering and incredibly at odds with our current practice. But it may be more than that. It may be that such a divorce would deeply undermine our sense of identity.

In the context of a discussion on strict liability, Tony Honré makes this point about identity. He says ‘outcome responsibility is being responsible for the good and harm we bring about by what we do. By allocating credit for the good outcomes of actions and discredit for bad ones, society imposes outcome responsibility’ (Honré, 1999: 14) and he claims that:

The best argument for outcome responsibility is surely that it is central to the identity and character of the agent. I contend that our responsibility for what we do and for its outcome is inseparable from our status as person. We are the people we are and have the character we have largely because the dealings in which our bodies and brains are involved, if in some aspect intentional, are attributed to us as the actions of persons with a continuing identity (Honré, 1999: 10).

He further goes on to argue that ascribing our outcomes to us is largely what gives us all our histories and our identities and that the price of having these is that we must be held responsible for the outcome of our actions (Honré, 1999: 29).

.

Honoré uses this idea, and some others, to argue for strict liability. I do not endorse strict liability nor do I think that we must be held outcome responsible for the outcomes of *all* our actions in order for us to retain our sense of identity. But I do think the outcomes of what we do are important to our sense of identity. If we were not held responsible for the outcomes to a significant extent it would be very alienating (and often dispiriting). Consequently I think that the credit and fortune we enjoy and discredit and costs we bear should, to a meaningful degree, arise from the outcomes of our actions.

Honoré's discussion is primarily concerned with bearing costs and benefits of a more material kind, but the point is pertinent to my discussion. If our outcomes are so important to us, then perhaps it is important that other types of responsibility (such as the degree of moral criticism we are liable to) are sensitive to them. Perhaps, the blaming-behaviour that others may have towards us must be sensitive to the outcomes that our actions end up having in order to preserve our sense of identity. The outcome for Matthew and Molly of killing or not killing a child will have an important impact on their sense of individual history. If our blaming-behaviour towards them did not reflect this then perhaps this would be alienating. The outcomes are not the *only* thing that matters when deciding how much to blame wrongdoers like Molly and Matthew, but I think that a satisfactory theory of punishment should probably allow us to subject Matthew to greater punishment than Molly.

Fortunately, using Scanlonian blame as the basis of our theory of punishment allows us to do this and therefore allows us to avoid being too radical and avoid undermining our sense of identity, even though it ties judgements of blameworthiness to an agent's moral

deliberations. In this section, I will explain how. Put simply, because it distinguishes between recognising that someone is blameworthy and actually blaming them, Scanlonian blame can use the idea of relationship impairment to permit a stronger blaming response to Matthew than Molly, even though it allows that Molly and Matthew are both equally blameworthy. It can thereby satisfactorily accommodate moral luck (or something very much like it).

I will first explain why Scanlon's account of moral blame can accommodate moral luck in the personal context⁷⁰ before explaining how this can be applied to punishment by the state.

Moral Luck in the Personal Context

As I discussed in chapter 3, blaming someone is more than judging that they are blameworthy, it also involves seeing your relationship with them as impaired. To summarise, blame involves:

1. A judgement: we judge that a person's deliberations are morally faulty.
2. A normative evaluation: we are judging the wrongdoer negatively for her faulty moral deliberations (we are not just noting it).
3. A conative component: judging that your relationship with the wrongdoer has been impaired.

⁷⁰ Here I am largely repeating Scanlon's own arguments (2008: 127, 148-149).

Judging that a person is blameworthy involves 1 and usually 2, but actually blaming her also requires 3. Because blaming someone includes 3, the extent to which it is appropriate to blame a person is not just influenced by how badly she has responded to moral reasons. It also depends on the wider significance of the event to the blamer's life. As Scanlon says:

As I interpret it, blame is not a *mere* evaluation but a revised understanding of our relations with a person, given what he or she had done. Blame is therefore a function not only of the gravity of a person's fault's but also of their significance for the agent's relations with the person who is doing the blaming (Scanlon, 2008: 150).

Again as noted in chapter 3, I can judge my boss to be blameworthy for cheating on his spouse. But, because my relationship with him is employer-employee, this has little significance for my life and consequently our relationship is not impaired. Therefore it would be odd and inappropriate for me to engage in blaming behaviour towards him. Here the significance of the blameworthy behaviour for me is affected by the relationship in which I stand with my boss. But other things can affect the significance of the blameworthy behaviour too.

Pertinent to the question of moral luck, is that one of these things is surely the actual outcome of the blameworthy behaviour. For instance, the significance of Matthew's actions in Stephen's parents' lives is of a totally different magnitude to the significance of Molly's actions in Sarah's parents' lives.⁷¹ The fact that Stephen is dead as a result of

⁷¹ Throughout I will talk about the significance of the death on the victim's parents' lives, rather than the significance of the death on the victim's life. I think that it does make sense to talk about the significance that the death has for the dead victim, but many do not (how can something be significant to someone who no longer thinks, feels, exists etc.?) so I want to avoid these references. Talking about the significance to the parents also means we can avoid bizarre hypotheticals concerning the impact of the victim's death on the post-mortem relationship between the victim and the killer. The reason I don't use an example with lesser harm is simply to stick with Scanlon's original example.

Matthew's actions is going to mean that interactions with Matthew will always have a certain kind of meaning for Stephen's parents. The fact that Molly's actions put Sarah at risk, will give Sarah's parents a reason to view their interactions with her as having a certain kind of meaning, but it will not be the same as the interactions between Stephen's parents and Matthew. Though they can negatively evaluate Molly's moral reasoning, it only has very minimal significance in their lives (see Scanlon, 2008: 149). This therefore means that the relationship impairment between Matthew and Stephen's parents will take on a different and more substantial character than the impairment between Molly and Sarah's and therefore will licence more blaming behaviour.

It is crucial to note that it is not *just* the fact that the outcome has a greater significance for Stephen's parents that means their relationship with Matthew is more impaired; it is the impact of the outcome on the significance of the *fault*. Scanlon brings this out nicely by comparing Matthew's case with a case of death by faultless driving:

Martin is disposed to be careful when he knows that the outcome of his actions might harm other people. He is a careful driver and has an up-to-date MOT test certificate etc.. One day, whilst he is driving carefully along, a small child, Sammy, runs out in front of the car and is killed.⁷²

It is not appropriate for Sammy's parents to blame Martin, because there is no faulty moral reasoning and it would not be plausible to say he is blameworthy. Nonetheless, the significance of Sammy's death would in general have a negative effect on the relationship between Sammy's parents and Martin.⁷³ (Blameworthy wrongdoing is not the only way to

⁷² I have modified the example (c.f. Scanlon, 2008: 148).

⁷³ The best outcome may actually be a situation where Sammy's parents and Martin grieve together and comfort one another, because both will be suffering – Sammy's parents from their loss, and

impair relationships: as I said in chapter 3, a relationship can be impaired, for example, if one party moves to Australia). The fact that Sammy died as a result of Martin's actions would always be there. Scanlon says that an 'objective stigma' (Scanlon, 2008: 127) would attach to Martin for Sammy's parents and it would involve a 'modification in [their] relations with' him (Scanlon, 2008: 148). This may have implications for the way in which they can treat him.

But the extent of the impairment between Stephen's parents and Matthew seems much greater than the impairment between Martin and Sammy's parents combined with the impairment between Molly and Sarah's parents (Scanlon, 2008: 150). The outcome of Matthew's action was partly down to his faulty moral reasoning and, as Scanlon says, 'the causal outcome of [Matthew's] action multiplies the significance of his fault' (Scanlon, 2008: 150). The multiplied fault in turn multiplies the impairment of the relationship – it leads to a much greater shift in the way that a relationship with Matthew can legitimately be viewed: In contrast to Martin, Scanlon says 'he is not only the person who killed a child, but also the person whose recklessness led to the child's death' (Scanlon, 2008: 149) and likewise, in contrast to Molly, he is not only the person who behaved recklessly, but also the person whose reckless behaviour killed a child. The fact that the significance of the fault is multiplied is, I think, important to providing an account that nicely satisfies people's intuitions on moral luck and it is this that helps justify the increased blaming response.

Martin because however blameless, he is going to have to live with the fact that he's killed a child. Nonetheless if Sammy's parents do not want to do this, this is certainly a forgivable flaw and one that most people probably possess. Thus we want to be able to explain implications for how Sammy's parents may treat Martin if they are unable to grieve with and comfort him.

So we can see that, in the personal context, Scanlonian blame can allow that one agent can correctly be subject to greater relationship impairment than another agent despite the fact that they have both displayed an equal level of faulty moral reasoning.⁷⁴ Thus it can deal fairly well with moral luck cases in the personal context. Now we need to consider whether this transfers satisfactorily to the context of state punishment.

Moral Luck in the Context of Punishment

In the personal context, we get the result that Matthew and Molly are equally blameworthy but that Matthew may be blamed more than Molly by the parents because the significance of the outcome multiplies the significance of the fault. We also get the result that, whilst Martin is not blameworthy, 'objective stigma' attaches to him with regards to Sammy's parents. But how does this apply to the context of punishment by the state? Would the significance of Matthew's fault be multiplied for people other than Stephen's parents? If not, would the state be allowed to punish him more than Molly? Would objective stigma attach to Martin with regards to the citizenry-at large? If so, would this mean his relationship with them was impaired in such a way as to allow the suspension of some of his citizenship rights? Clearly this would not be an acceptable outcome. I shall consider Martin's case first, then Molly and Matthew's.

Objective stigma does not attach to Martin with regards to the citizenry-at-large. The child's death is not of appropriate significance to licence the citizenry-at-large to view

⁷⁴ It is also an advantage of Scanlon's account of blame that it can allow for the idea that something similar to blame exists in cases like Martin's.

their relationship with Martin as impaired. Recall from chapter 4 that for an action to licence citizens to view their *citizenship* relationship with an actor as impaired, the action must (i) fall inside the sphere of actions where citizens are thought to act together and (ii) it must undermine citizen solidarity. With regards to Martin's actions, only the first condition is met, so his co-citizens do not have the standing to blame him. Composing and upholding traffic regulations clearly fall within the realm of actions where citizens act together. But, since Martin has followed the traffic regulations, his action does not undermine citizen solidarity. Therefore even if there may be some objective stigma attached to Martin, this does not undermine the co-citizen relationship, so the citizenry-at-large is not licensed to alter the way they interact with Matthew *qua* co-citizen. Thus the state is not licensed to punish him.

Both Matthew's and Molly's actions licence the citizenry-at large to blame them. This is because they (i) fall inside the sphere of actions where citizens act together and (ii) they undermine citizen solidarity. This is because, as I already said, citizens act together when composing and upholding traffic regulations and reckless car driving is the sort of thing that should be illegal. This illegal reckless car driving clearly undermines citizen solidarity. This therefore impairs the co-citizen relationship and licences some form of punishment. However, if Sammy's death is not relevant to the citizenry-at-large's relationship with Martin, must the blame-based expressivist be committed to the view that equal punishment should be given to Molly and Matthew? I will now suggest two possible reasons why not.

First, recall that the significance of Matthew's *fault* is multiplied by the outcome of his actions. It is plausible to think that the significance of his fault is legitimately viewed as multiplied by everybody, not just by Stephen's parents. This sounds plausible: it seems right that everyone can recognise that Matthew's fault has greater significance than Molly's. Even though Stephen's death would have nowhere near the same significance for the citizenry-at-large as it would for his parents, the fact that a child has died (in a context which is the business of the citizenry-at-large) still has significance. If this is right then it seems that the citizenry-at-large can actually blame Matthew more than Molly. If this is true, then they can punish Molly and Matthew to different degrees.

I think there is some plausibility in this argument. However, it is not the main justificatory route I wish to travel as it may suffer from a version of the 'wrong-victim' problem I identified in chapter 4. Recall that it is not usually satisfying to say that we should base our justification of wrongdoers on the idea that they've wronged the citizenry-at-large. It is better that our justification centres on the wrong done to the *victim*. I gave the following example: the suggestion that the state has the standing to punish rapists because rapists wronged the citizenry-at-large is quite strange. If someone asked me 'what gives you the standing to blame the man who raped Susie?' I don't think the right answer is 'because he's wronged *me*'. Rather the answer needs to be about the wrong done to Susie. Likewise (though to a lesser extent), if we are going to punish Matthew more than Molly, I think it would be unsatisfying to claim that this was because of the greater significance of Stephen's death to *all* the citizens. It would be better if we could centre our response on the significance to Stephen's parents.

This brings me to the second reason why blame-based expressivism does not require equal punishment of Matthew and Molly. Recall that the citizenry-at-large is licensed to blame *on behalf* of the victims. As I argued in chapter 4, when citizens claim the standing to blame wrongdoers, they are not claiming that they are somehow the victims of that wrong. They are claiming that their relationship of solidarity and representation with the victim makes the blameworthy wrong that has befallen her into their business and they can blame vicariously on her behalf. We can say the same thing about the multiplied significance of the fault caused by the outcome of Matthew's action. We have already said that Matthew's action is one that is the legitimate business of the citizenry-at-large. So the citizenry-at-large's relationship of solidarity and representation with Stephen's parents licences them to *vicariously* take up the multiplied significance of the fault and view their relationship with him as more impaired than their relationship with Molly. Therefore blame-based expressivism allows us to take the outcomes of a criminal's actions into account when deciding how much to punish them (whilst avoiding the 'wrong victim problem').

This section sought to answer the question of whether accounts of blame that tie a person's blameworthiness to their deliberations must be committed to denying that moral luck exists and, if so, whether my account of state punishment must be committed to the idea that individuals like Matthew and Molly should receive equal punishment. An account of punishment committed to these things would be highly radical and may undermine our sense of identity. However, I have argued that Scanlonian blame can allow for moral luck because it takes into account the significance of fault and that this idea can be transferred to the context of punishment by the state, so individuals like Molly and Matthew need not be punished equally. I now move on to look at whether blame-based expressivism can

avoid the objection that it must licence the state to punish non-harmful, trivial and private blameworthiness.

I. Trivial, Private, Unrelated And Non-Harmful Blameworthiness and the While Life's Virtue Objection

In chapters 3 and 4, I argued that the fact that a citizen is blameworthy is, in part, what should justify her punishment. But one can be morally blameworthy for *trivial* wrongs, such as stealing a fiver, or *private* wrongs, such as cheating on one's spouse. It is also possible to be morally blameworthy for actions that do not risk or cause harm (henceforth '*non-harmful wrongs*'). Hate speech is a prime example. Though hate speech is appalling, it doesn't always risk or cause harm. Yet it is certainly possible to be morally blameworthy for it. It is even possible to be blameworthy without even performing a relevant action (henceforth '*blameworthy non-acts*'). For example, if I witness a brutal assault and am secretly cheering for the aggressor inside my head, then I am certainly blameworthy. But I am not blameworthy for the *action* because I am not agent responsible for it. It is also possible to be blameworthy for an action that is unrelated to the one now in question (henceforth '*unrelated-acts*'). Imagine, for example, that I am framed for crime Y now, but that I committed an equivalent crime, X, 10 years ago and was never caught. I am blameworthy for X but not for Y, but does this matter to how others may now react to me? Many find it profoundly counterintuitive to think that the state should punish trivial, private, unrelated and non-harmful acts and non-acts. But, if it is the blameworthiness that justifies the punishment, then why should these things not be punished? Why, indeed,

should the state not punish and reward a citizen according to the moral virtue she has accrued over the course of her whole life (henceforth the '*whole life's virtue objection*')?

Similar questions to these are much discussed in the philosophical literature. The objections relating to trivial and private wrongs are often pressed⁷⁵ at 'rights-forfeiture views' of punishment which state, roughly, that by performing a wrong act, a citizen loses her right against hard treatment. So, if performing a wrong act makes a citizen lose her right, why doesn't performing a trivial or private wrong act also have this result? If people push these criticisms against rights forfeiture accounts, they would certainly push them against my account.⁷⁶ We can be blameworthy for trivial and private wrongs so my account is open to this line of objection.

The objection relating to non-harmful acts and non-acts has been pressed on accounts of the justification of self-defence which argue that an aggressor needs to be morally blameworthy for the threat that they pose to another's life in order to be fully liable to be killed by them in self-defence. The objection is that, if it's the aggressor's blameworthiness that matters, then a victim should be able to use a non-threatening but blameworthy bystander to defend herself from the threat (perhaps she could use her as a shield). However unlikely this scenario is, it is thought to undermine the credibility of the idea that blameworthiness is necessary for full liability to self-defence: the bystander hasn't posed a risk of harm, it doesn't seem right that she should be liable (see e.g. McMahan, 2005: 393). It would be no giant leap, then, for someone to apply this criticism

⁷⁵They have been pressed, for example, by David Boonin (2008: 115-117)

⁷⁶ I conclude that my account seems a natural target from conversations with my supervisor.

to my justification of punishment. We might imagine an objector to blame-based expressivism wondering whether it would licence the state to punish blameworthy actions that are not harmful or even blameworthy thoughts with no accompanying harmful actions.

The objection concerning unrelated acts has been pressed against rights forfeiture views (by e.g. Lippke 2001a: 80-81; Quinn 1985: 332–33). It is often called the *relatedness problem*. Quinn asks us to imagine that the state imprisons a bunch of political dissidents, but in the process mistakenly imprisons one who is not guilty of that offense but is guilty of an unknown about past offense that would have rendered him liable to an equivalent punishment. His intuition is that it is wrong for the state to punish him for dissidence rather than for the crime he actually committed. But, he believes that the rights forfeiture view cannot account for this: if he has forfeited his right against the punishment, why does it matter which crime he is punished for? One can easily imagine then that this objection could be pressed against blame-based expressivism since (i) it draws on elements of the rights forfeiture tradition and (ii) one can be blameworthy for unrelated acts.

The whole-life's virtue objection has been pressed against desert-based retributive views by Victor Tadros (Tadros, 2011: 69-70). Desert-based retributivism is the view that the state should punish offenders because it is good for wrongdoers to suffer. Tadros's question is that, if offenders should suffer because of their wrongdoing, then how can desert-based retributivists resist the conclusion that the state should try to make people's level of well being track their aggregate level of good and bad deeds throughout their entire lives? It seems inappropriate for the state to do this, so this poses a problem for desert-based retributivists. Does a similar problem apply to blame-based expressivism? We

could see why one might think so: if blameworthy wrongdoing impairs your relationship with the state, does praiseworthy action enhance it? If so, why should the state not make citizens' entitlements track their aggregate level of praiseworthy and blameworthy acts throughout their entire lives?

I will now explain how these lines of objection can be overcome.

Trivial Blameworthy Acts

Some people find the idea of the state punishing very trivial instances of law breaking counter-intuitive. In order for this intuition to be plausible we have to stipulate that we really do mean *trivial* law breaking (such as smoking a joint, underage drinking, stealing ten pounds, etc.) rather than significant law breaking that does not risk serious harm (like tax fraud, electoral fraud, etc). I share this intuition. However, I am not convinced that we need there to be a principled reason against punishing trivial wrongs in order to satisfy our intuitions. Rather, I think our intuitions can be satisfied by the idea that punishing trivial law breaking just isn't worth it (c.f. Wellman, 2012: 389). Blame-based expressivism can account for this intuition. The argument mirrors the one given in the above discussion about moral luck. I assume that there is a moral obligation to obey the law and thus that trivially breaking the law shows faulty moral reasoning (but see section II). Since trivial law breaking is law breaking, it is the sort of fault that is the legitimate concern of all citizens. However, it seems to be true by definition that trivial law breaking has no great significance for anybody's life and is probably also not *that* blameworthy. Therefore, since

relationship impairment is determined by the blameworthiness of an action and the significance of it to people's lives, trivial wrongs will not produce very much relationship impairment. If trivial wrongs hardly produce any relationship impairment, then it is hardly worth punishing them.⁷⁷

Private Blameworthy Acts

Blame-based expressivism is not vulnerable to the objection that it must allow the state to punish individuals for private blameworthy actions. This is because, as I discussed in chapters 3 and 4, judging that a person is blameworthy for X does not give you the *standing* to blame her. To re-use an example, suppose my boss had an extra-marital affair and, further assume, that in this instance I can legitimately judge him to be blameworthy for this (he has failed to give pertinent moral reasons sufficient weight in his reasoning). It is far from clear that I stand in a relationship to him that makes this affair my legitimate concern and therefore far from clear that I have the standing to view my relationship with him as impaired and act accordingly. If I decided to call him out on his behaviour or cease to engage in normal employee-employer civilities with him, we might reasonably conclude that I do not really have the standing to do this as it's none of my business or because I retain employee-employer duties to him that I should uphold (c.f. Scanlon, 2008: chpt 4; Coates and Tognazzini, 2016; Radzik 2011: 582).

⁷⁷ Christopher Wellman endorses the claim that 'no [non-consequentialist] theory of punishment' is able to give an explanation of why the state should not punish trivial acts without referring to arguments about utility because 'there is no principled reason that *de minimis* crimes may not permissibly be punished' (2012: 389). Perhaps blame-based expressivism is an exception to this claim.

So if we want to know whether a person has the standing to blame, we must establish that the blameworthy behaviour in question is one that is the legitimate business of the person wishing to engage in blaming behaviour. Likewise if we want to know whether the state has the standing to punish a citizen for a particular piece of blameworthy behaviour, we must first establish if that behaviour is the legitimate concern of the citizenry-at-large. It is by no means automatic that that it is – it needs to be *explained* why citizens are licensed to view their relationships with wrongdoers as impaired when they are not the victims. Once we recognise that the judgement that a person is blameworthy does not give everyone the standing to blame her, we can see that there is no presumption that blame-based expressivism justifies punishing private blameworthy acts – rather the burden of justification lies more in attempting to explain why the state had standing to blame at all.

Providing this justification was one of the main tasks I set out to accomplish in chapter 4. There I argued that the wrong done to the victim was the citizenry-at-large's legitimate concern if it (i) falls inside the sphere of actions where citizens are thought to act together and (ii) it undermines citizen solidarity. I also argued that it seemed minimally controversial to say that law breaking meets these conditions and therefore that the citizenry-at-large usually has the standing to blame when someone breaks the law. This means citizens would be protected from punishment for private blameworthy acts. It is true, however, that I did not say much about what the content of the law should be. I did not give an argument as to whether or not, for example, adultery could permissibly be made a crime. Someone might thus object that if blameworthiness is so important to my account of punishment, wouldn't it also be important to my account of what counts as a

crime? However, this need not be the case. My thesis relies on the idea that when someone commits a crime she does something blameworthy and that this blameworthiness licences the citizenry-at-large to view its relationship with her as impaired. However, this is not the same as claiming that the reason why certain actions should be illegal is that they are blameworthy. Therefore, there is no reason to suppose a society adopting blame-based expressivism as a theory of punishment would be particularly prone to an erosion of the private sphere which claimed private blameworthy acts as the legitimate business of all citizens.

In any case, I intend my theory to apply to liberal democracies. Thus, I intend it to apply to a society which affirms individual freedom and recognises that there must be some room for a private life. No palatable conception of liberal citizenship will conclude that every aspect of a person's life is the state's business. It will not conclude that every area of a citizen's life should be thought of as an area of joint-action which links to co-citizen solidarity. Liberal conceptions of citizenship will leave room for private blameworthy acts and, importantly, their criminal law will reflect this.⁷⁸ Thus, since blame-based expressivism is intended for a state which affirms these values, it will be constrained by them. Saying that blame-based expressivism will be constrained in this way is not *ad hoc*: it is recognition that a theory of punishment must be situated within a theory of the proper scope of state action and citizenship. It is important that individuals have scope for private wrongs. In the case of blame-based expressivism, there is no reason to assume that allowing the state to pass judgement on the blameworthiness of its citizens when their actions fall into the remit of legitimate state business means that it must punish them for private blameworthy acts.

⁷⁸ Duff makes this point too (2001: 64).

In dismissing the objections that blame-based expressivist should be committed to punishing private wrongs, I have allowed the idea that external principles may be appealed to. By this I mean, that when the state is determining the content of its criminal law, it must appeal to some moral principle other than those which most directly underlie blame-based expressivism. One might object to this appeal to external principles as a kind of hand-washing exercise. However, I think this an unrealistic and unappealing objection. It is unappealing because we live in a world with many values and so surely want our state to reflect this. It is unrealistic because attempts to reduce everything to a single principle must, I am sure, sneak in other principles by the back door. I believe that many theories of punishment need to appeal to external principles to exclude trivial and private wrongs from punishment. For example, even if you think the only legitimate reason for state action is harm prevention and thus think that the general justifying aim of punishment is to prevent harm, you are not safe from this worry. Someone could say to you, ‘If you care about preventing harm, why not care about private harms?’ I’m not claiming a proponent of this view cannot find a way to respond, but she will have to bring in other considerations (either at this point or in coming up with her definition of what counts as harm). Thus the need to bring in external theories to limit the state’s scope is not a reason to be put off blame-based expressivism *in particular*. Here, to some extent, I echo the sentiments of Christopher Wellman (2012: 391).⁷⁹ He notes that if there are good arguments why the state should not punish certain types of wrongs, it is open for any theorist (who affirms value pluralism) to invoke them.

⁷⁹ I say ‘to some extent’ because Wellman is defending rights-forfeiture theories of punishment but his arguments can be applied to protect blame-based expressivism.

Non-Harmful Acts and Non-Acts

It is controversial to claim that the state may punish acts which don't cause or pose a significant risk of harm. However, my account need not be committed to punishing non-harmful acts and non-acts.

First let's take the case of non-harmful non-acts. Recall the example of the witness to a brutal assault who is secretly cheering for the aggressor inside her head. Consider also a person who only refrains from acting out her murderous fantasies because she is scared of getting caught. My account of punishment does not require the state to punish these people. This is for two reasons. The first echoes the arguments made in relation to moral luck. The second echoes the arguments made in relation to private wrongs. The first reason is that a person's fantasies are not plausibly the sort of thing that should be considered an area for joint action and no palatable conception of citizenship would try to criminalise fantasies (if indeed it were possible). Therefore the state will not have the standing to blame citizens here. The second reason is that non-acts like this seem unlikely to have any major significance for their 'victims'. Therefore there is no reason to think that the blameworthy people in these examples should be punished.

Now let's consider non-harmful blameworthy acts. Let's take hate speech. I believe that hate speech certainly often has great significance for its victims, even if it is not physically harmful. This then would put a presumption in favour of blame-based expressivism punishing blameworthy hate speech. I do not take this a decisive objection against blame-

based expressivism – I am not that opposed to punishing hate speech.⁸⁰ However, in line with the arguments about private wrongs, *if* the value of free speech and belief is so important, then this presumption can be overridden.

Unrelated Acts

It is possible to be blameworthy for an action that is unrelated to the one the state now seeks to punish you for. Imagine, for example, that I am framed for crime Y now, but that I committed an equivalent crime, X, 10 years ago and was never caught. Let's assume that I am blameworthy for X, I must therefore have damaged my relationship with my co-citizens. Does this now mean that, according to blame-based expressivism, the state may permissibly punish me for crime Y even though it would be punishing me for the wrong crime?

I have two responses. The first goes as follows: In our example, it is true that by committing crime X I have damaged my relationship with my co-citizens. However, recall from chapter 3 that blame-based expressivism is an expressive theory in the sense that the positive reason to punish comes from the idea that it is good for the state to express criticism of wrongdoing and that withdrawing citizenship rights is a permissible way for the state to express this criticism. In our example, if the state withdraws citizenship rights from me it cannot be doing that to express criticism of me for committing crime X because it does not know I have committed this crime. Instead it would be doing this to express

⁸⁰ Not least because I do not believe that accounts of harm should be limited to material harm and thus that hate speech is usually harmful. But I have responded to this objection simply because it is a controversial issue and it is an advantage of a theory to be able to be flexible towards it.

criticism of me for committing crime Y. But, Blame-based expressivism is not a matter of keeping an overall ledger of wrongs committed and punishment inflicted, but of responding appropriately to a particular wrongful act. It seems very uncontroversial to assume that one does not render oneself liable to criticism for things that one has not done even if one is liable to criticism for other things that one has done. It also seems uncontroversial to assume that there is no (non-instrumental) positive value in criticising people for things they have not done. Therefore I am not liable to criticism for crime Y and criticising me for doing it would not be valuable, so the positive reason to punish – that the state should express criticism of wrongdoers – does not apply here.

I now give my second response. According to blame-based expressivism, the other positive reason to punish wrongdoers is to maintain the state-victim relationship. This positive reason also does not apply in our example. In our example a positive reason the state has to punish me is to maintain its relationship with the victim of crime X. If it punishes me for crime Y, it does not do that. Therefore, according to blame-based expressivism the state would have no positive reason to punish me for committing crime Y. This further shields blame-based expressivism from the objection that it would permit the state to punish unrelated acts.

The Whole Life's Virtue Objection

Tadros presses the whole-life's virtue objection upon desert-based retributivism. Tadros makes the following points about desert-based retributivism: First, if moral wrongdoing is

the appropriate type of wrongdoing for which people deserve to suffer then surely they should suffer for all the wrongs they commit throughout the course of their life. Second, if wrongdoers deserve to suffer then do-gooders should deserve an increase in well-being. Third, if a citizen commits a mixture of wrongs and goods, her well being should match this. I shall now explain whether each of these points can be used as a basis of an objection to blame-based expressivism.

- a. Punishing wrongdoers for all of their wrongs.

The reason Tadros's first point bites against desert-based retributivism is that, in taking moral wrongdoing as the appropriate basis for deserved suffering, it seems to be silent on the question of which wrongs we deserve to suffer for and which wrongs *the state* may make us suffer for. Desert-based retributivists have been thought to find it hard to non-arbitrarily demarcate which wrongs the state may make us suffer for and are therefore thought to be committed to the claim that the state may make us suffer for all wrongs we commit. This would have the implication that the state may punish an individual for all the wrongs he committed throughout his whole life which would be implausible because it implies that the state may punish him for private, trivial, non-harmful, and unrelated wrongs. However, though desert-based retributivism may have this implication, blame-based expressivism does not. As I argued above, blame-based expressivism can explain why the state may not punish citizens for private, trivial, non-harmful, and unrelated acts. Also, as I mentioned in the previous section, blame-based expressivism is not concerned with keeping a moral ledger – rather it responds to particular acts. Therefore, it only has the implication that the state may punish an individual for all the *criminal* wrongs that he

committed throughout his whole life. But clearly this is not an implausible implication - it is what should be the case.

b. Rewarding praiseworthy acts.

Desert-based retributivism states that wrongdoers deserve to suffer. When we cause a person to suffer we reduce his well-being. Therefore desert-based retributivism makes a recommendation for what should happen to the well-being of wrongdoers. Therefore, the argument goes, desert-based retributivists are committed to the reverse idea that do-gooders deserve an increase in their well-being. This seems an inappropriate goal for the state to pursue. Is blame-based expressivism committed to something similar?

Blame-based expressivism does not make recommendations about well-being (even though what it recommends is likely to have an impact on well-being). So blame-based expressivism would not be committed to the idea that the state should seek to raise the level of well-being of do-gooders. Rather what it might entail is that citizens who carry out praiseworthy acts enhance their relationship with the state and thus should be entitled to special privileges. This would be wholly implausible if it meant that every and any type of praiseworthy act would give a citizen such privileges, but it would not mean this. As we saw above and in chapter 4, the state may only blame citizens for wrongs that it has the standing to blame them for – the state-citizen relationship is only damaged if the wrongs committed are the legitimate business of the state. Therefore the same would follow for praiseworthy acts – citizens would not be entitled to special privileges for all or any

praiseworthy acts, only those that are appropriately considered the state's business. I do not propose to set out exactly which praiseworthy acts generate gratitude or should count as the business of the state, but I think we could categorise it broadly as acts of public service. When rendered this way, I do not think the proposition that citizens should be awarded special privileges for praiseworthy acts is so implausible.

According to Scanlon, the counterpart of his conception of blame as relationship impairment is not praise, it is gratitude. Praiseworthy acts therefore generate the reciprocal duties that we normally associated with gratitude:

Gratitude is not just a positive emotion but also an awareness that one's relationship with a person has been altered by some action or attitude on that person's part.

Sincerely felt, this entails having the reciprocal attitudes that this change in relationship makes appropriate. Typically this involves a readiness to respond in kind – for example, a greater readiness to help a person who has gone out of her way to help you (Scanlon, 2008: 151).

The idea that we should have gratitude towards those who dedicate themselves to public service and that this might give rise to special reciprocal duties is highly intuitive. In 1997, the percentage of ex-military personnel in London's homeless populations was 22. Isn't that shocking? Surely one of the reasons why it is so shocking is that military personnel have sacrificed a lot with the aim of making our state safe. We should be grateful for this and we owe them more than to let them become homeless. Of course, we shouldn't let any of our co-citizens become homeless, but it seems particularly bad in the case of those who have served our country. As well as being highly intuitive, the idea that we owe others

duties of gratitude and reciprocity also enjoys considerable philosophical support (for gratitude see (Berger, 1975); for reciprocity see e.g. Sangiovanni, 2007).

So why might one think that the idea that we owe citizens who perform relevant praiseworthy acts (like public service) duties of gratitude is an unattractive implication of my view? One reason could be because of what it might imply about praiseworthy citizens' statuses. Throughout the thesis I have said that when a citizen commits a wrong, this undermines their relationship with other citizens and at times I have said that this therefore undermines their *status* as co-citizens. Would this therefore mean that those who perform relevant praiseworthy acts enhance their relationship with their co-citizens and thereby enhance their citizenship status? If so, would that create a status hierarchy between citizens who devote themselves to public service and those who do not? If it does, would that be an unacceptable implication?

I believe that a sort of hierarchy does emerge but that it is not an objectionable one. To explain why, I am going to consider two ideas. First, recall from chapter 4 that the claim that citizens should have equal membership of their political community is distinct from the claim that all people have equal moral status in virtue of their humanity. Thus if we grant that a praiseworthy citizen has a higher status, this does not mean that we are saying she has a higher status in the sense that she has more moral worth than an ordinary citizen (someone who follows the laws, but is not in the habit of performing acts of great public service). Second, I want to consider more carefully the sense in which a praiseworthy citizen could be thought to have higher status. I will begin by considering an example concerning friendship. Let's suppose that in my friendship group I have an outstanding

friend, a steady friend, and a bad friend. My outstanding friend is the best friend, not in the sense that she's the closest, but in the sense that she best follows the virtues of friendship, such as, always remembering significant dates, never cancelling on me because of an exciting turn in her love life or because she's tired, spending long hours deliberating with me about a not very important option I'm considering etc. My steady friend competently adheres to the requirements of friendship, but is not quite as virtuous as my outstanding friend. By contrast my bad friend is not meeting the requirements of friendship, he's always late or cancelling and has started being rude and argumentative. Let's consider the implications of this on their statuses. My bad friend is compromising his status as my friend. He has started to put himself outside the friendship circle. My outstanding friend is a better friend than my steady friend in the senses described and her outstanding behaviour as a friend gives me reason to be grateful to her. But, assuming that the other factors which can influence friendship are fairly constant between the two friends, it is perfectly plausible to say that both of them hold equally solid places in my friendship circle. Thus although my outstanding friend is a better friend in the sense that she upholds the virtues of friendship better and therefore has a higher status in that regard, it is perfectly plausible to say that both my steady and outstanding friend have equal status in the sense of being my friends.

If we transfer the analogy to the context of citizenship, we can see that something similar holds there. It seems to me that doing public service can make someone a better citizen in the sense that she is adhering to the virtues of citizenship better than ordinary citizens. Thus in this sense she has a higher status than the other citizens. But I do not think her praiseworthy actions make her more of citizen than ordinary citizens. It is perfectly plausible to think that both praiseworthy and ordinary citizens hold equally firm in their

status as citizens. This is because, unlike blameworthy citizens, ordinary citizens have not done anything to jeopardise their citizenship status - the state has set down a set of rules that it expects its citizens to follow, and if they do not break these rules it is not plausible to think they undermine their citizenship status. If this is right, then although there is a type of hierarchy, it is not objectionable – it is hardly objectionable to say that some people are more virtuous citizens than others.

What could be objectionable, however, is the idea that those devoted to public service get special privileges. Whether it is or not will turn on what those special privileges turn out to be. I cannot conduct an extensive review of this question, but let me consider a few obvious candidates and explain why I do not think that they pose a problem. Honours, I take it, are not a cause for concern – so long as they are awarded fairly for genuine important public services (and do not give people places in a powerful and unelected chamber of government). Nor, it seems to me, is monetary reward troubling. For example, many people think that public sector workers should receive generous pensions. There may be a few reasons for this (such as that people believe many public sector workers to be underpaid), but it is plausible to assume that gratitude for their hard work in serving the public is one of them. We might think that because blame-based expressivism can permit us to withdraw the right to vote from wrongdoers, it would also imply that citizens who perform praiseworthy acts are entitled to greater political privileges, such as votes with enhanced weighting. This would be a troubling implication but it does not follow from blame-based expressivism. As I argued in chapter 4, the reason blameworthy citizens lose their rights to vote is that, through their actions, they are putting themselves outside the group of citizens. But, as I just argued, praiseworthy and ordinary citizens are both firm

members of this group. Thus it is unclear why we should think they would be entitled to votes with enhanced weighting.

Another reason why one might think that the implication that we owe praiseworthy citizens a duty of gratitude is problematic is as follows: I have argued that wrongdoing citizens lose their citizenship *rights*, and this thus might imply that praiseworthy citizens gain extra *rights* that we treat them with gratitude. But, the objection goes, though we can have reason to express gratitude to those who perform such acts, it is problematic to think that these acts ground *rights*. Perhaps one thinks that giving someone what they are entitled to doesn't really seem to be plausibly thought of as expressing gratitude – one might think that for something to be a true expression of gratitude it must come freely from the giver rather than be something (like a right) that could in principle be demanded by the receiver (Wellman, 1999); or perhaps one might think that just because I have moral reason to give someone a thing of mine out of gratitude, this does not show that he is entitled to it (McDermott, 2004; Wellman 2012). However, this is not really pertinent to this discussion. If we cannot have rights that people show us gratitude, then blame-based expressivism cannot generate the conclusion that citizens can gain them – it only generates the conclusion that co-citizens have gratitude based reasons to award special privileges to those that perform the relevant praiseworthy acts or that praiseworthy citizens gain special claims or privileges, rather than rights.

- c. Citizens that commit both praiseworthy and blameworthy acts.

As I said above, the third point concerning the whole life's virtue objection that Tadros makes against desert-based retributivism is that if a citizen commits a mixture of wrongs and goods, her overall well being should match this. For desert-based retributivism this could be a problem because, as Tadros argues, it might imply for example that we should encourage a miserable do-gooder to do wrong so that her well-being better matches her virtue. It could also be a problem because it might imply that if a citizen is generally very good and has a normal level of well-being, we should not punish her for her wrongdoing if she commits a crime. As we know, blame-based expressivism does not make recommendations about what wrongdoer's well-being levels should be, but might a similar objection to this apply. Might it be the case that if a citizen commits a mixture of relevant praiseworthy and blameworthy actions, she could avoid punishment of her wrongs?

This would be the case if the aggregate rewards a citizen was owed cancelled out the aggregate punishment. However, this is not the most logical way of understanding the situation. First, it is not clear that rewards and punishments do cancel each other out. If I am owed gratitude and a generous pension because of my services to the public that does not mean that I am immune from blame for committing a crime. This is because these are separate things. It is therefore much more logical to think that I should be punished for my wrongs and separately rewarded for my praiseworthy actions. Therefore I do not see that blame-based expressivism is committed to the counter-intuitive implication that a citizen can avoid punishment by committing relevant praiseworthy actions.

In this section I have considered the objections that blame-based expressivism would require the state to punish trivial, private, unrelated and non-harmful blameworthy acts and non-acts or make a citizen's entitlements match the aggregate level of praise and blameworthy acts over her whole life. I argued that it does not. It is not required to punish trivial blameworthy acts of law breaking. The significance of such acts is small and so is the blame that attaches to them. Therefore the size of the relationship impairment would be small because this depends on the person's blameworthiness and the significance of their action's outcome. I argued that blame-based expressivism does not require the state to punish private wrongs because it does not have the standing to do so. Blame-based expressivism is perfectly compatible with the recognising individual freedom and that there must be room for citizens to have a private life. I argued that blame-based expressivism does not require the state to punish non-harmful acts and non-acts: both because it can lack the standing to do so and because such acts and non-acts have little significance. I also argued that blame-based expressivism is not vulnerable to the unrelated acts and whole life's virtue objections. This is largely because blame-based expressivism is not about keeping an overall ledger of wrongs committed and punishment inflicted, but of responding appropriately to a particular wrongful act. I now move on to consider objections relating to the state having to assess a citizen's moral reasoning.

III. State Assessment of its Citizens' Moral Reasoning

Blame-based expressivism requires the state to make a judgement that a citizen is blameworthy before it punishes her. When we judge that a person is morally blameworthy for an action, we judge that she is agent responsible for it and that the way she has authored it shows that her moral deliberations about why to pursue that action were *morally faulty*. If judging that a person is blameworthy involves judging that her moral deliberations about why to take an action were morally faulty, then blame-based expressivism would require the state to judge and punish a person on the basis of her moral reasoning.

Many people believe that the quality of a citizen's moral reasoning is none of the state's business and, thus, that it would be objectionable for the state to require citizens to reveal it and then judge them on its basis. Matt Matravers and Victor Tadros believe this:

The modern democratic state is one in which the writ of the state is not thought to extend into the mind, motivation, and 'private life' of its citizens (Matravers 2000: 93).

There is a limit to what we can expect people to divulge about their moral convictions in court. An investigation into a person's deeper moral views would often wrongly invade their privacy (Tadros, 2011: 164).

This sounds similar to the private wrongs objection but it is not just a different version of that objection. The private wrongs objection says that some *wrongs* are private (like adultery and lying to friends). This objection says that *the quality of a citizen's moral*

reasoning is not the state's business. Perhaps this is because of a right to privacy or perhaps it is because of the value of free thought and belief or perhaps it is because of the value of state neutrality.

The objection is also not that the state may never inquire as to what is going on inside the mind of a citizen. For example, we want to be able to excuse or exonerate or exempt people from punishment on the grounds of diminished responsibility. Therefore we must allow the state to assess whether a citizen is capable of being held agent responsible either in general or at the time when a crime was committed. This then would require the state to, amongst other things, assess whether a citizen is capable of seeing certain reasons for action as reasons for action and, if she is, whether she is capable of acting on them or whether, for example, some mental process is preventing this etc.

The objection also cannot be that the state may never inquire into what a citizen's reasons for actions are. It's hard to make sense of existing practices of punishment unless you allow the courts to enquire into motives. There are many offences in which the motive seems an essential part of establishing which crime has been committed. For example, consider trying to decide if a medical professional has committed murder or medical manslaughter. The state may need to inquire into whether this person administered a lethal dose of morphine with the intention of killing or whether she administered it with the intention of relieving suffering even though she foresaw that it would bring about death. Many theorists also believe that we need to know a citizen's surface level motive to understand whether she can be excused (on some theorists' interpretations e.g. (Gardner,

2007: 103)) or justified (on other theorists' interpretations e.g. (Tadros, 2007: 268)) for breaking a law.⁸¹

I think the objection is to the idea of the state trying to make citizens reveal the deeper moral beliefs that underlie their motives and actions and then judging them on the basis of these beliefs. To assess the precise level of a person's blameworthiness, we do need to know her deeper moral beliefs. Let me give an example to illustrate why. Imagine that Harry and Sally desecrate a Jewish grave. Sally does it because she believes that Jewish people are the scum of the Earth, whereas Harry only does it in order to impress Sally. We could ask Harry and Sally about their immediate motivation and Sally might say something like 'I wanted to bring the dead person's relatives down a peg or two' and Harry might say 'I wanted Sally to think I was brave and rebellious'. From this we can assess that Sally is more blameworthy than Harry. But to really judge the precise degree of blameworthiness we need to go a bit further. We need to ask Sally 'Why did you want this?' and she might reply 'they're evil people who sponge off society'. We might ask Harry 'But what about the feelings of the relatives?' and maybe he would say 'I feel bad about that, but I was stupid and wanted to impress Sally' thus implying that, at a deeper level in his belief structure, he doesn't have an objectionable view about the worth of Jewish people. We can now judge Sally is even more blameworthy than Harry than we

⁸¹ Some theorists do not think that the state inquiring as to what motivated a citizen is pertinent to establishing excuse/justification or the difference between medical manslaughter and murder. They say that what matters is whether conditions exist that would justify/excuse a reasonable person or that would mean a reasonable person could be convicted of medical manslaughter instead of murder. However, I think these theorists are probably moved by considerations of fairness (if two people perform the same action under the same set of external conditions, they should be treated the same) or pragmatism (it is probably easier to find out what a reasonable person could claim motivated them, than what actually motivated a person). I doubt they would think it was morally objectionable for the state to try to find out what a person's motives are *just as such*.

could before. Therefore we do at least sometimes need to know a person's deeper moral beliefs to make a precise evaluation of their blameworthiness.⁸²

So how can the blame-based expressivist respond to this? First we should note that it is relatively common to think that the state may occasionally inquire into a citizen's deeper moral beliefs. It needs to do this, for example, if it wants to treat civil disobedients and conscientious objectors differently from other lawbreakers. In the case of both civil disobedience and conscientious objection, the state needs to know that the lawbreakers hold a particular moral belief with regards to why they broke the law and to this extent needs to know about their deeper moral beliefs. Thus I don't see how any state that wants to allow different treatment of conscientious objectors can avoid requiring them to divulge that they have a specific moral belief that sits in contrast to the law they wish to break. The state also needs to assess a citizen's deeper motives if it wants to convict them of terrorism: according to the UK's Terrorism Act (2000), Terrorism involves a threat or harm 'for the purpose of advancing a political, religious, or ideological cause'. However, these cases seem to be the exception rather than the rule and I accept that we would not want the state to make inquiries into citizens' deeper moral beliefs as a matter of routine. If it did it may plausibly threaten privacy, free belief, and state neutrality. If blame-based expressivism would require scrutiny of this deeper kind as a matter of routine, then this would be an important objection against it. But would it? There are two arguments why we may think it might.

⁸² Some might prefer to say badness here but I put this to one side.

Argument 1 is that we might think blame-based expressivism implies that a person's punishment should track the precise degree of her blameworthiness for breaking a specific law: the more blameworthy a person is for breaking a law, the more she impairs her relationship with the victim and therefore the more she should be punished. If so then, since we need to know a person's deeper moral beliefs to know the precise level of her blameworthiness, the state would need to routinely examine wrongdoers' deeper moral beliefs. I think there is some truth to argument 1. I think blame-based expressivism does imply that punishment should track a wrongdoer's blameworthiness for breaking a particular law. But I think it only implies this in the sense that it creates a presumption in favour of the idea and I think this presumption can easily be defeated. If it is important that the state does not routinely assess citizen's deeper moral beliefs for reasons of privacy, neutrality, or free thought, then a blame-based expressivist can appeal to these ideas so long as she can still give a satisfactory explanation of how the state can establish that the wrongdoer is blameworthy. I explain how it can do these things in my response to argument 2.

Argument 2 is that we might think that we need to establish what a person's deeper moral beliefs are in order to be able to establish that she is blameworthy or in order to be able to make appropriate assessments about the degree to which a wrongdoer is blameworthy. This argument is false. We need to know about a person's deeper moral beliefs to know her *precise* degree of blameworthiness. But we do not, in general, need to know about a person's deeper moral beliefs to be justified in making appropriately differentiated attributions of blameworthiness. We do not in general need to look into a person's deeper moral beliefs to be able to make differentiated attributions of blameworthiness because, in general, the more serious the wrong a person commits, the more blameworthy she will be.

In general, a person will be more blameworthy for rape or murder than she will be for petty theft.

We also do not, in general, need to know about a person's deeper moral beliefs to be justified in presuming simply that she is blameworthy for breaking the law. This is because we can be justified in presuming that a person is blameworthy simply for failing to take a pertinent moral reason against acting into account. If we go back to our example of Harry and Sally, although we needed to look at deeper moral beliefs to determine that Sally is more blameworthy than Harry, we did not need to look at them to be pretty confident that both were blameworthy. Assuming they were both agent responsible for their actions and had no other excuse, it is enough to know that they failed to take two pertinent moral reasons against acting (that it is wrong to desecrate graves and that it is wrong to break the law) as a sufficient moral reason not to act. Likewise, imagine that Timmy intentionally kills Toby against Toby's will. Timmy is not being coerced, he is not high, Toby is not threatening him etc. We do not need to know whether Timmy has killed for love, money, or malice in order to know that he has not taken 'do not kill' as a sufficient moral reason against acting. This is enough to create a strong presumption in favour of him being blameworthy. We do not need to know about his deeper moral beliefs. I think this is applicable to most cases where breaking the law is obviously wrong. Thus I think that we do not usually need to look into person's deeper moral beliefs to be justified in presuming that she is blameworthy for breaking a law which prohibits an obviously wrong action.

But what about cases where the morality of the action is subject to reasonable disagreement? An example of such a case might be where a husband has been moved by

compassion to carry out the assisted suicide of his terminally ill wife. In such cases, it is not so clear that breaking the law makes a person morally blameworthy. In such cases, whether a person is blameworthy or not might turn on what her moral beliefs are. Thus the state might need to find out what a citizen's deeper moral beliefs are to find out if she is blameworthy in such cases. In order to be excused of blame for breaking the law, one must have (i) a reasonable and genuine moral belief that the act itself is morally permissible (ii) a reasonable and genuine moral belief that one has sufficient reason to disobey the law and (iii) one must have been motivated to break the law by these beliefs. What this means though is that people are only really going to have a chance at being excused because of their moral beliefs if they are reasonably objecting to the law for moral reasons. This category of actions is broader than conscientious objection in the sense that conscientious objection occurs when someone fails to uphold the law because she believes she is morally *prohibited* from following it. But it is similar to it and I think that it is reasonable to allow the state to delve into a citizen's deeper moral reasons in cases like these. Thus I think that, so long as the law breaker is agent-responsible for her crime, it is usually acceptable for the state to presume that she is blameworthy unless she can reasonably claim otherwise. Therefore I think the claim made in argument 2 – that we need to establish what a person's deeper moral beliefs are before we can establish that she is blameworthy – is false. Thus the state can usually establish that a wrongdoer is blameworthy for committing a crime without an investigation into her deeper moral beliefs.

I shall now consider an objection to the idea that we do not need to know about deeper moral beliefs to establish blameworthiness. Some philosophers think that people can be excused of blame for wrongful acts if they have mistaken moral beliefs. If our mistaken

moral beliefs can excuse us of blameworthiness then a blame-based expressivist system may sometimes need to investigate a citizen's deeper moral beliefs. The mere fact that a person believes she is doing the right thing, morally speaking, is not sufficient to nullify her blameworthiness. Imagine that a person unlawfully kills another because her religion commits her to the view that adulterers should be stoned to death, or a person kills someone because she wants to 'liberate' his organs for the rest of society to use. The fact that these people believe they are doing the right thing might make them slightly less blameworthy than someone who murders sadistically. But they are still blameworthy. The reason for this is that the beliefs that one should stone adulterers to death or liberate people's organs are obviously substantively morally mistaken and, absent some further excusing explanation, the fact that the agents do not recognise this shows that their moral reasoning is faulty.

The most plausible case where we think a person may be excused of false moral beliefs is when she has an epistemic excuse. We might think, for example, that Wallace could be significantly excused for having racist moral beliefs if she grew up in 1920's Atlanta. However, even where a person does have an excuse for an objectionable belief, this does not mean that she is excused of blame for all the actions that arise from that belief: they will not significantly excuse her for committing obvious wrongs. She would not be excused of blame at all for committing serious crimes against people (like lynching) even if she could be excused for having racist beliefs. So in these kinds of cases, her deeper moral reasoning would not need to be assessed. All we would need to know was that she didn't give due consideration to the moral reason she had against acting. We might, however, think that Wallace could be significantly excused of blame for some of the actions motivated by her racist beliefs that are not quite so obviously wrong. For example,

she *might* be significantly excused for some cases of employment discrimination and therefore the state may elect to treat her differently from someone who did not have her excusing background. However, if the state wants to treat Wallace differently – and I’m really not sure that it should - then it needs to make assumptions about deeper moral beliefs in this case. But it can base that assumption on the context that Wallace has come from rather than from an investigation into her particular deep beliefs. Therefore, even if the state is assessing her deeper moral beliefs, it does not need to be a problematic invasion of her privacy, nor does it seem to undermine the values of free thought and belief. It is also worth noting that cases where we can excuse a person for having unreasonable moral beliefs would be really quite rare in a liberal democracy. They may crop up, for example, in relation to immigrants from very different cultural contexts or in the case where, for example, octogenarians are being tried from crimes committed in their twenties.

Conclusion

In this chapter I considered the objections that blame-based expressivism would not be able to deal with moral luck; trivial, private, unrelated or non-harmful blameworthiness; the whole life’s virtue objection; and that it would not be able to prevent the state from making objectionable inquiries into its citizen’s deeper moral beliefs. In considering these objections I hope to have covered some of the more immediate qualms that one might have on hearing the proposal that a citizen’s blameworthiness forms one of the main reasons that justifies state punishment and I hope to have shown that blame-based expressivism can respond to them in a satisfactory manner. The idea that there is a distinction between having the standing to judge a person blameworthy and having the standing to actually

engage in blaming behaviour towards her has played an important role in enabling my responses. It allows us to see why blame-based expressivism is consistent with the idea that citizens should have a private life and some degree of free thought, which in turn can explain why we need not punish citizens for private and non-harmful acts as well as why we should limit the extent to which a state probe's into a citizen's deeper moral reasoning when it is trying to establish her blameworthiness. I believe that a failure to appreciate this distinction is probably what lies at the root of some of the most uncomfortable initial qualms around the proposal that punishment should be a response to a citizen's blameworthiness. This chapter is therefore important because I suspect these qualms may lead to philosophers dismissing blame-based theories of punishment out-of-hand. But, in fact, if we subject the qualms to deeper exploration, as I have done in this chapter, we can see that they need not trouble us. This then safely opens the door for a route of theorising about punishment that is of great potential utility and has many intuitive features.

Concluding Remarks

I. Short Summary

In this thesis I have sought to defend an account of the justification of punishment that draws upon both the expressive and rights forfeiture traditions. Using the Scanlonian conception of blame gives us the resources to start to specify which rights are lost when a person is guilty of criminal wrongdoing: they forfeit the rights that arise out of their special relationship with their co-citizens (their citizenship rights). This is not, as some other retributive and rights forfeiture views are, based on any mysterious idea of cancelling out the wrong done, but rather on the more commonsense idea that our relationships are partly constituted by what we do and therefore the special rights that they ground can also be affected by what we do. Our basic rights, however, are not grounded on special relationships and thus are not affected by changes to our special relationships. Citizenship rights include rights to move freely around the public spaces of one's state and the right to vote. Basic rights include rights to have our basic physical, social, and autonomy-based needs are met. Blame-based expressivism therefore respects the crucial insight from rights forfeiture views – that we must explain why wrongdoers forfeit their rights against punishment. It also starts to provide a positive account of which rights are forfeited that sets plausible limits on what the state may do in the name of punishment. This is in contrast to most other rights forfeiture views which either do not provide a positive account of which rights are forfeited or rely on the implausible strict reciprocity view that the rights forfeited are the rights the offender has violated.

Relationship impairment then gives the state a *prima facie* permission to punish wrongdoers because it explains why doing so does not violate wrongdoers' rights. But it does not explain whether punishment is all-things-considered permissible. It does not give us a positive reason to punish. We need a positive reason because we need to show why it is worth spending state resources on punishment rather than on other things. Here I pointed to the positive value of expressing moral criticism of wrongdoers. In my view, it is important that the state expresses this criticism for its own sake, to maintain solidarity and a sense of moral community amongst the citizenry-at-large, and, in cases with specific victims, to maintain the relationship between the state and the victims by showing the victims that it cares about the wrong done to them. It is worth allocating state resources to these goals even though it means that these resources cannot be spent on other goods such as healthcare. I want the state I live in to pursue a plurality of goods and I think that it is worth taking, say, an increased risk of not being able to have an injury or disease treated quite as well as it otherwise might have been, in order to pursue the goals of showing victims that we care and maintaining solidarity amongst the citizenry-at-large.

Expressive theories have not typically been able to explain why the moral criticism of wrongdoers must, or even may, permissibly take the form of punishment. This is because there are many ways of expressing moral criticism. Many expressivists make the contextual claim that verbal criticism is not enough to express our moral criticism. I take this claim seriously, but we still need to show why moral criticism must or may take the form of punishment, rather than some other method that goes beyond the merely verbal. If we use the idea of blame as relationship impairment, we can meet this challenge as we can now explain why punishment is analogous to the types of blaming behaviours that we show to our other special relations. When we direct blaming behaviour towards a friend,

we may do this by suspending the special claims that her friendship would normally allow her to make on us. If the state withdraws citizenship rights from wrongdoers, this is simply an analogous kind of behaviour, even though it can have more serious implications. This represents a step forward for expressive views which have typically been unable to either explain why punishment comes within the permissible range of blaming behaviours or link it directly to other intuitively permissible forms of blaming.

I hope my arguments are convincing or at least that they might be useful or interesting because they provide a different interpretation of expressive and rights forfeiture views. I also hope some of the negative arguments I have given could have use in helping to prompt further ideas and debate. In particular, I hope some of the arguments I have given show that it is important to reflect on the form of punishment that a particular theory permits. This is not often given much attention by philosophers, but the implications one's theory has for the form of punishment it permits will affect the overall plausibility of the theory.

My account also raises some unanswered questions and implications that are worth considering. I turn briefly to these now and outline the avenues of future possible research that they open.

II. Ideas for Further Research and Implications for Other Research Topics

How should we think of basic rights and human rights?

Basic rights, recall, are the rights that need to be protected so that we can have a minimally decent life. They are grounded in our humanity and not on the basis of any special relationship that we are in. The idea that wrongdoers retain basic rights is very attractive and is endorsed by, for example, the United Nations Basic Principles for the Treatment of Prisoners (1990) which states that ‘all prisoners shall be treated with the respect due to their inherent dignity and value as human beings’. However, this idea has implications for how we should think of basic rights. I believe these implications have hitherto gone unnoticed, but are worthy of further research.

Some philosophers hold a generous interpretation of what it takes to have a minimally decent life and, accordingly, argue for an expansive set of basic rights to ensure such a life. However, this does not sit easily with the idea that we can imprison people in such a way that they maintain their basic rights: it seems unlikely that life in prison, even a relatively open prison, can be consistent with a *very generous* conception of what constitutes a minimally decent life. This poses a trilemma for someone who is committed to an expansive list of basic rights. It means that they have to accept one of the following three ideas:

- (i) that we may not imprison wrongdoers
- (ii) that wrongdoers should not keep their basic rights
- (iii) that punishment is a permissible rights infringement⁸³

⁸³ One could claim that wrongdoers do not lose their basic rights, that prison is not consistent with having rights protected, but it is nonetheless permissible to imprison wrongdoers even though it will infringe their basic rights.

The last possibility might do for occasional punishments. When a lot is at stake, it is sometimes permissible to infringe a right – for example, it might be permissible to infringe your property right by breaking into your empty house if it is necessary to do so to shelter from a very dangerous storm. So it might be permissible, for example, to infringe the rights of a very dangerous offender through a disproportionately long prison sentence, when we know that releasing her would be very likely to result in multiple violent crimes. But this must only be for very exceptional cases: we do not want our standard justification of punishment to imply that it is a permissible rights infringement. When we permissibly infringe someone's rights, we typically think we owe them some form of compensation. But, providing we have stated within the bounds of proportionality, it would be counterintuitive to claim that (as standard practice) we owe wrongdoers compensation for imprisoning them. When we permissibly infringe a person's rights, we also typically think that we have wronged them in some way, albeit justifiably so. But, except in very exceptional cases, it would be counterintuitive to claim that punishment wrongs wrongdoers.

The first possibility may be right, but as I have said before this is a radical proposal and the implications need to be thought through. If we seriously think that the first possibility is correct then this means that, as a matter of great urgency, we need to start reconceptualising how the state should punish. If one disagrees with me that the first possibility is incorrect, then I hope that my work has value in the sense that it highlights the need for this re-conceptualisation.

What of (ii)? Perhaps, one might think, wrongdoers do lose their basic rights (or some of them at least). A desert-based retributivist, for example, might be quite happy with this

thought. But the argument that wrongdoers forfeit their basic rights it is not an easy one to make. As some of my argument in chapters 1 and 2 should illustrate, it is difficult to explain *why* wrongdoers lose basic rights when they commit crimes. Even if one manages to do this, one would then need to establish a rationale for determining which basic rights were lost (one would have to rule out the possibility that it is permissible to punish through removing basic rights against physical harm, for example). Perhaps this is an achievable task, but it would not be an easy one. Therefore theorists who want to argue for an expansive conception of basic rights cannot simply assume that rejecting (ii) provides the answer to the trilemma.

There are also implications here for how we might want to think about human rights. Some theorists working in this domain like to keep their list of human rights fairly minimal, whereas others are inclined to make it more expansive. The theorists who like to keep their lists fairly minimal often do so because they believe we only have human rights to a minimally decent life and thus only basic rights count as human rights (Miller, 2012; Wellman, 2016). Some of those who argue for a more expansive set, however, also believe that we only have human rights to a minimally decent life - they just have a more demanding account of what one needs for a minimally decent life and thus argue for a more expansive set of basic rights (e.g. Oberman, 2016). Other theorists who opt for a more expansive set of human rights think we can have human rights to more than just a minimally decent life. They thus believe that more than just our basic rights can be human rights and more than just our basic needs can ground our human rights (Cole, 2015). They are pluralists about what can ground human rights. Because of this pluralism, this latter group of theorists can plausibly claim that citizenship rights – rights that one has only in

virtue of membership of a state - can be included on the list of human rights.⁸⁴ It is not as plausible for theorists who argue that only our basic rights are human rights to claim this. This is because our basic rights are not supposed to be contingent on our membership in a group. Thus, for those who prefer an expansive set of human rights, the trilemma I have identified might be a reason to favour a pluralist account of the grounds of human rights. This is because it would allow us to claim that the right to a reasonable degree of free movement around a state is a human (citizenship) right (which I assume is something that we want to do), whilst also preserving a category of human rights (basic rights) that cannot be lost by wrongdoers.

The cost of this approach is that it might mean rejecting the ambition that all our human rights should be recognisable as valuable by all people, whatever culture they are from (for supporters of this ambition see e.g. Miller, 2012; Wellman, 2016). This is because the idea that free movement around a territory is a citizenship right is partly based on the idea that, as equal members of the political community, citizens are entitled to move freely around the public spaces of a state. This in turn relies on the liberal conception of citizenship as equal membership of a political community, which may not be acceptable in all cultures. Perhaps the cost of rejecting this ambition is higher than the cost of rejecting one of the three claims posed in the trilemma. But, either way, I think this is an issue worth exploring in future research.

Having the capability to distinguish between human rights based on citizenship (e.g. the right to move freely around the territory) and human rights based on our humanity (e.g.

⁸⁴ This is conditional on them accepting that human rights are rights that all people possess, rather than, as some philosophers claim, as rights that all people possess against *all* other people. (All people can possess citizenship rights because all people should be members of a political community).

basic rights) might also be helpful for the newly emerging literature on immigration. For example, Christopher Wellman is currently puzzling over the apparent asymmetry between the right to move freely within a territory (currently recognised as a human right) and the right to move freely between territories (currently not recognised as a human right). He is drawn to the ‘highly revisionist’ conclusion that accepting a restriction on free movement between territories means that states may restrict citizens’ free movement within territories without violating their human rights (Wellman, 2016: 89-90, quote at 90). However, if the right to free movement is a human citizenship right, rather than a basic right, then this might help to explain the apparent asymmetry.⁸⁵

Completing My Account of Which Rights are Citizenship Rights

Through blame-bases expressivism I have developed a rationale for explaining which rights may be suspended during punishment (citizenship rights) and which rights may not (basic and other protected rights). I have also tried to supply a rationale for determining what count as basic rights (they are the rights that need to be protected if we are to have minimally decent lives). But at the moment I am lacking a clear rationale for determining what should count as citizenship rights vs. what should count as non-basic kinds of protected rights (recall that a ‘protected right’ is one that is not liable to be removed in punishment according to blame-based expressivism). For example, though I think the right to due process is a protected right, I am lacking a rationale (rather than just *ad hoc* reasons) to explain why. Without this rationale, my theory of punishment is incomplete because we

⁸⁵ I doubt this solution would be acceptable to Wellman as he is drawn to the view that only our basic rights are human rights and he has a fairly restricted view of what counts as a minimally decent life (Wellman, 2016: 89).

do not have a full account of which rights can and cannot be suspended according to blame-based expressivism. Developing such a rationale would therefore be an important line of future research for a blame-based expressivist.

One implication that I would need to consider is that, according to many theorists, citizenship gives rise to rights to greater redistribution of resources within states than between states (for example of three different views of this kind see Blake; 2001; Sangiovanni, 2007; Miller, 2007). For example, one might think that we have duties to give everyone in the world the resources needed to meet their basic rights, but that we only have a duty to co-citizens to ensure they have enough resources to participate in the economic sphere of our state on equal terms. This could take the form of less well-off citizens receiving benefit payments from the state. If it is true (as I think it is) that citizenship gives rise to rights to greater redistribution of resources, then, in principle, this right can be suspend from wrongdoers. Of course, the state could leave wrongdoers without enough resources to meet their basic rights, but it would still be a troubling implication. This would not really be that relevant to wrongdoers whilst they are in prison, as they are to some extent cut off from society at large and do not really have many opportunities to spend money. However, it would be relevant to those on Home Detention Curfew and doing Community Service.

Can blame-based expressivism avoid this implication? I think so. This is because we have reason to argue that it is not all-things-considered permissible for the state to withdraw the right to more extensive redistribution of resources. One reason could be that the suspension of citizenship status is not permanent and, when we punish, we need to make sure that the wrongdoer can come back into the fold easily enough. If we make it too

difficult for them to participate equally in economic terms this could make their return unnecessarily difficult. Another important reason is that many wrongdoers come from economically disadvantaged groups and, as I shall consider below, may even have been the victims of economic injustice at the hands of the state. Insofar as being in a disadvantaged group is due to injustice and/or factors beyond the wrongdoer's control, it would clearly be wrong and/or unfair to heap more economic disadvantage on them. This suggests that it would be impermissible for the state to punish wrongdoers by withdrawing economic benefits.

III. To What Extent can Blame-based Expressivism be applied to the Real World?

Disadvantaged Wrongdoers

I believe that the UK is a reasonably just society, but this is not to say that no injustices are perpetrated. Thousands of children in the UK live in relative poverty and I think it is fair to say that not all UK citizens are given enough opportunity to participate in the social, political, and economic spheres of the state on equal terms. We can plausibly consider these individuals as the victims of injustice at the hands of the state. This is because the state has a duty to ensure that its citizens do not find themselves disadvantaged in these ways. However, as I just said, a great many people who commit crimes come from these disadvantaged groups. Assuming that there is a relevant causal connection between poverty and crime, we might think the state is partly complicit in the crimes of disadvantaged wrongdoers. In chapter 4, I described how someone can lack the standing to blame someone else when the wrong committed is none of her business. However, it is also commonly recognised that people can lack the standing to blame wrongdoers when

they are partly complicit in her wrong (see Cohen, 2006; Duff, 2010). It would usually be wrong of me to unleash blaming behaviours on you for a crime we just committed together. So, if the state is part complicit in the wrongdoing of the poor, why would it have the standing to punish them? One response here would be to say that blame-based expressivism is only intended to apply to a society that was free from economic injustice. But I do not want to give this response. Given that many of our criminal wrongdoers do come from disadvantaged groups, it is dull at best and complacent at worst to sidestep the issue in this way.

I cannot give a full response here but I will sketch some of the issues that we need to consider and ways that we might try to proceed. We should start by noting a few things about the way the state is complicit in the crimes of disadvantaged wrongdoers. The state does not usually compel disadvantaged citizens to commit crimes – there are usually other options available. The state does not intend or encourage disadvantaged citizens to commit crimes. The state does not even intend to make disadvantaged citizens disadvantaged. At worst, it is responsible for negligently putting disadvantaged citizens in a position where it makes it more difficult for them to avoid being responsible for committing criminal acts. At best, it is non-negligently responsible for putting them in such a position - ridding the UK of economic injustice is not exactly an easy task. There are empirical and political difficulties that stand in the way of such goals. So the state may have been doing what we could reasonably expect and yet still fail to rid itself of economic injustice. This means that the state's standing to blame is not as damaged as it would be if it compelled, intended or encouraged wrongdoers to commit crimes, but, it does not leave the state's standing intact either.

How should this problem be resolved? May the state punish disadvantaged wrongdoers or not? Tadros proposes that this problem is, in a sense, irresolvable – whatever we do, we will have reason to regret it. If we go ahead and punish, we have reason to regret this since, he argues, it means that the state has failed to hold itself responsible for its part in the crime. However, if we do not punish, we also have reason to regret our decision because the state would not give the (often also disadvantaged) victim the acknowledgement she's due. It may reduce the impact of deterrence and consequently undermine society's security (Tadros, 2009).

Duff is more hopeful. He poses the first step of a solution: the state can have the standing to blame if it is willing to answer for its own failings too (Duff, 2010). If the state were complicit in disadvantaged citizens' crimes in the sense that it had compelled, intended or encouraged them to commit them, then perhaps this idea might be more difficult to sustain. Perhaps in some of these cases, the state would bear a much greater responsibility for the crime than the wrongdoer and this perhaps might mean that the state does not get to unilaterally decide that both parties will be called to account. However, since the state is only complicit in the sense that it has made it more difficult for disadvantaged wrongdoers to avoid crime, I think Duff's idea is attractive.

How should we proceed with it? Duff is unsure but tentatively suggests 'the development of fora in which unjustly disadvantaged citizens can pursue their legitimate grievances' (Duff, 2010: 139). I think this is a promising idea in principle, but may suffer in application. Public fora of this kind might be shaming since, in the real world, many people are ashamed of their poverty. Maybe a better way forward would be to ensure that the state builds extra support and opportunities into the parole period for disadvantaged

wrongdoers. I will not explore this possibility further here but we can note that the question of whether the state has the standing to punish the disadvantaged has received very little philosophical attention. Deeper research into this question seems important for those who want to consider the permissibility of punishment in our current society.

Feasibility

In this thesis I have sought to defend the use of prisons. I have not, of course, sought to defend prisons as most of them currently are. Most of them currently violate basic rights and are therefore unjust. Thus, though my thesis is not abolitionist, it is reformist. But what guide to action does it give to those who want to engage in prison reform today?

I defend the idea that prisons should give inmates many more opportunities to pursue their conception of the good, participate in public life, and spend time with loved ones than they currently have. I also argued that these opportunities need to be ‘adequately delivered’. We may be able to find some examples of the sorts of prisons I have argued for (perhaps, for example, in Finland or perhaps in some of England’s ‘open’ prisons). But running large scale institutions is difficult, especially on a relatively constrained budget and when the prisoners you are managing will be going through stress, do not want to be there, and may have the tendency to be somewhat unreasonable. I fear that my view of what a prison should look like would be viewed as hopelessly naïve by some of those who actually work in the prison services. They might think that setting up such institutions simply isn’t achievable in today’s climate.

If we follow Geoffrey Brennan and Nicholas Southwood in defining a feasible outcome as one that an agent is reasonably likely to be able to achieve if she or it tries to achieve it (Brennan and Southwood, 2007: 5-10), then we might start to query the feasibility of the types of prisons I have argued for. Thus if we believe that 'we ought to X' implies 'X is feasible', then, maybe the claim we ought to set up the type of prisons I have argued for is false. If this is the case, then what implications would it have for what we should or may do about prison reform right now? Maybe our governments should just try their best to set up the sorts of institutions I argue for or maybe they should try to set up institutions that are as close to them as *is* feasible. But this is not always the right answer to feasibility problems and I think it is unlikely to be in this case. This is because, though these efforts may result in prisons that are better than many of the prisons we currently have, they are still likely to violate inmates' basic rights.

Maybe we need to investigate further possibilities. Perhaps it is feasible to set up prisons that do not violate basic rights if the state makes the institutions smaller than we are used to and with larger staff: inmate ratios. This would make it easier for staff to have more detailed knowledge of the inmates so they can appropriately differentiate the opportunities offered and build the positive relationships needed for effective implementation. This is likely to make prisons more expensive and so would mean that the state would probably have to scale back our use of prisons and combine it with greater use of other types of punishment, for example Home Detention Curfew.

We are also faced with question of what we, as individuals or as prison reform collectives, should or may do if the state does not comply with feasible options to make the sort of prison system that would respect basic rights. Should we simply just keep trying to

persuade them to comply? Should we just take any opportunity that might bring about a reduction in the violation of prisoners rights, even though we know that the sorts of institutions we campaign for will still violate some of prisoners' basic rights? What if doing this might ultimately make it harder to persuade the state not to violate wrongdoers' basic rights at all?

There is a growing literature on what 'feasibility' means (see Brennan and Southwood, 2007; Gilabert and Lawford-Smith, 2012); what conditions need to be met for the actions of *collectives* (such as states) to be feasible (Lawford-Smith, 2012); what collective duties mean for the individuals within a collective (Lawford-Smith, 2012); what we should do when we cannot achieve our ideal and whether we even need to know what our ideal is in order to know what we should do (Sen, 2006; Simmons, 2010; Stemplowska, 2008); and what we should do when others fail to comply with their duties (Miller, 2011; Stemplowska, 2016).

This literature has been applied to questions of global justice. But to my knowledge this literature has not been widely applied to the domestic context in general (what is it feasible to expect states to do for their own citizens?) and I do not think it has been applied at all to the question of prison reform in particular. But it would be good to make headway in this topic. Prison reform is an urgent area of moral action in the real world: if I am right that prisoners do not forfeit basic rights, then thousands of prisoners' basic rights are wrongly being violated inside the borders of liberal democracies *right now*. As such, it is pressing to find out whether it is feasible to set up the sort of penal institutions I have argued for. If so, can we do it on a large scale or are we permitted only to have a small prison population? If not, what should or may we do instead?

Bibliography

Adamson, H. 2011. 'What Does Blame Do to Relationships?' *Praxis* 3(1): 1756-1019

Altman, A. and C. H. Wellman. 2009. *A Liberal Theory of International Justice*. Oxford: Oxford University Press

Arpaly, N. 2006. *Merit, Meaning, and Human Bondage: An Essay on Free Will*. Princeton: Princeton University Press

Barry, Christian, and Nicholas Southwood. 2011. 'What Is Special About Human Rights?' *Ethics and International Affairs*, (25): 1-15

Berger, F. 1975. 'Gratitude', *Ethics*, 85 (4): 298–309

Blake, Michael. 2001. 'Distributive Justice, State Coercion, and Autonomy', *Philosophy & Public Affairs*, 30 (3): 257-296.

Boonin, D. 2008. *The Problem of Punishment*. New York: Cambridge University Press

Brownlee, Kimberley. 2012. 'What are the Duties in the Duty View?' *Jerusalem Review of Legal Studies*, 5 (1): 62-74.

Brownlee, Kimberley. 2012. 'Social Deprivation and Criminal Justice' in François Tanguay-Renaud & James Stribopoulos eds. *Rethinking Criminal Law Theory*. Oxford: Hart Publishing

Buchanan, A. 2004. *Justice, Legitimacy, and Self-Determination*. Oxford: Oxford University Press

Caney, S. 2005. *Justice Beyond Borders*. Oxford: Oxford University Press

Carens, Joseph. 2016. 'In Defence of Birthright Citizenship' in Sarah Fine and Leah Ypi eds. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press

Carens, J. H. 1987. 'Aliens and Citizens: The Case for Open Borders', *The Review of Politics*, 49 (2): 251–73

Coates, D. Justin and Neal Tognazzini. 2013. 'The Contours of Blame' in D. Justin Coates and Neal A. Tognazzini eds. *Blame: Its Nature and Norms*. Oxford: Oxford University Press

Cohen, G. 2006. 'Casting the first stone: Who can, and who can't, condemn the terrorists?' *Royal Institute of Philosophy Supplement*, 58: 113-136

Cohen, Joshua. 2006. 'Is There A Human Right To Democracy?' in Christine Synowich eds. *The Egalitarian Conscience: Essays in Honour of G. A. Cohen*. Oxford: Oxford University Press

Crook, F. 2006. 'Close Down Women's Prisons', *The Guardian*, Available at: <https://www.theguardian.com/commentisfree/2006/aug/02/closedownwomensprisons>.

Darwall, S. 2012, 'Bipolar Obligation' in R. Shafer–Landau eds. *Oxford Studies in Metaethics*. Oxford: Oxford University Press

Dolinko, D. 1991. 'Some Thoughts About Retributivism', *Ethics* 101, (3): 537-559.

Doob, A. N. and Cheryl Marie Webster. 2003. 'Sentence Severity and Crime: Accepting the Null Hypothesis', *Crime and Justice*, 30: 143-195

Duff, Antony. 2013. 'Legal Punishment', in Edward N. Zalta ed. [online] *The Stanford Encyclopedia of Philosophy*, Available at <http://plato.stanford.edu/archives/sum2013/entries/legal-punishment/>

Duff, A. 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press

Duff, R.A. 2010. 'Blame, Moral Standing and the Legitimacy of the Criminal Trial' *Ratio*, 23 (2): 123-140

Durlauf, Steven N. and Daniel S. Nagin. 2010. 'The Deterrent Effect of Imprisonment' in Philip Cook, Jens Ludwig, and Justin Mccrary eds. *Controlling Crime: Strategies and Tradeoffs*. Chicago: University of Chicago Press

Durlauf, Steven N. and Daniel S. Nagin. 2011. 'Imprisonment and Crime: Can Both Be Reduced?' *Criminology & Public Policy*, 10 (1): 9-12

Fabre, C. 2000. *Social Rights Under the Constitution: Government and the Decent Life*. Oxford: Oxford University Press

Falls, M. M. 1987. 'Retribution, Reciprocity, and Respect for Persons', *Law and Philosophy*, 6: 25-51

Feinberg, J. 1970. 'The Expressive Function of Punishment', in J Feinberg eds. *Doing and Deserving*. Princeton: Princeton University Press:

Gilabert, P. and Lawford-Smith, H. 2012. 'Political feasibility: A conceptual exploration', *Political Studies*, 60(4): 809-825

Gilabet, P. 2009. 'The Feasibility of Basic Socioeconomic Human Rights. A Conceptual Exploration', *The Philosophical Quarterly* 59 (237): 559-581

Goldman, Alan H. 1979. 'The paradox of punishment', *Philosophy & Public Affairs*, (9)1: 42-58

Griffin, James. 2001 'First Steps In An Account Of Human Rights', *European Journal of Philosophy*, 9(3): 306-327

- Griffin, J. 2008. *On Human Rights*. Oxford: Oxford University Press
- Hampton, Jean. 1992. 'Correcting Harms versus Righting Wrongs: The Goal of Retribution', *UCLA Law Review*, 39: 1659–1702
- Hart, H. L. A. 1963. *Law, Liberty and Morality*. New York: Random House
- Hanna, N. 2008. 'Say what? A critique of Expressive Retributivism', *Law and Philosophy*, 27(2): 123-150
- Von Hirsch, A. 1993. *Censure and Sanctions*. Oxford: Oxford University Press
- Von Hirsch, A. and Anthony E. Bottoms, Elizabeth Burney, and P-O. Wikstrom. 1999. *Criminal Deterrence and Sentence Severity: An Analysis of Recent Research*. Oxford: Hart Publishing
- Honoré, Tony. 1999. *Responsibility and Fault*. Oxford: Hart Publishing
- Husak, D. 2013. 'Retributivism In Extremis', *Law and Philosophy*, 32(1): 3-31
- Kershnar, S. 2000. 'A Defence of Retributivism', *International Journal of Applied Philosophy*, 14(1): 97-117
- Kramer, M.H., 2011. *The Ethics of Capital Punishment: A Philosophical Investigation of Evil and its Consequences*. Oxford University Press.
- Lawford-Smith, H. 2012. 'The feasibility of collectives' actions', *Australasian Journal of Philosophy*, 90 (3): 453-467
- Lazar, S. 2010. 'A Liberal Defence of (Some) Duties to Compatriots', *Journal of Applied Philosophy*, 27 (3): 246-257

Liebling, Alison and Ben Crewe. 2012. 'Prison Life, Penal Power, and Prison Effects' in Mike Maguire, Rod Morgan, and Robert Reiner eds. *The Oxford Handbook of Criminal Law*. Oxford: Oxford University Press

Lindquist, C.H., 2000. 'Social Integration and Mental Well-Being Among Jail Inmates', *Sociological Forum* 15 (3, pp. 431-455). Kluwer Academic Publishers-Plenum Publishers.

Lippke, R. 2001a. 'Criminal Offenders and Rights Forfeiture', *Journal of Social Philosophy*, 32: 78–89

Lippke, R. 2001b. 'The Disenfranchisement of Felons', *Law and Philosophy*, 20: 553–580

Lippke, R. 2007. *Rethinking Imprisonment*. Oxford: Oxford University Press

Loucks, Nancy. 2000. *Prison Rules: A Working Guide*. London: Prison Reform Trust

Marshall, Sandra E. and R. Anthony Duff. 1998. 'Criminalization and sharing wrongs', *Canadian Journal of Law and Jurisprudence*, 11: 7-22

Mason, A. 1997. 'Special obligations to compatriots', *Ethics* 107(3): 427–47

Mason, A. 2012. *Living Together As Equals: The Demands of Citizenship*. Oxford: Oxford University Press

Mason, M. 2011. 'Blame: Taking it Seriously', *Philosophy and Phenomenological Research*, LXXXIII (2): 473-481

Matravers, M. 2000. *Justice and Punishment: The Rationale of Coercion*. Oxford: Oxford University Press

McDermott, D. 2001. 'The permissibility of punishment', *Law and Philosophy*, 20(4): 403-432

- McDermott, D. 2004. 'Fair Play Obligations', *Political Studies*, 52 (2): 216-232
- McMahan, J. 2009. *Killing in War*. Oxford University Press
- Miller, D. 1997. 'Equality and Justice' *Ratio*, 10(3): 222-237
- Miller, D. 2002. 'Group Rights, Human Rights and Citizenship', *European Journal of Philosophy*, 10(2): 178-195
- Miller, D. 2007. *National Responsibility and Global Justice*. Oxford: Oxford University Press
- Miller, D. 2009. 'Justice and Boundaries', *Politics, Philosophy & Economics*, 8(3): 291-309
- Miller, D. 2012. 'Grounding Human Rights', *Critical review of International Social and Political Philosophy*, 15(4): 407-427
- Miller, David. 2016. 'Is There a Human Right to Immigrate?' in Sarah Fine and Leah Ypi eds. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press
- Ministry of Justice. 2016. *Prison population figures: 2016*. London: Ministry of Justice.
- Moore, M.S., 1993. 'Justifying Retributivism', *Israel Law Review*, 27(1-2): 15-4
- Morris, C. W. 1991, 'Punishment and Loss of Moral Standing', *Canadian Journal of Philosophy*, 21: 53-79
- Nagel, T. 2005. 'The Problem of Global Justice', *Philosophy & Public Affairs*, 33(2): 113-147
- Narayan, U. 1993. 'Appropriate Responses and Preventive Benefits: Justifying Censure and Hard Treatment in Legal Punishment', *Oxford Journal of Legal Studies*, 13: 166-82

Nickel, J. W. 2005. 'Poverty and Rights', *The Philosophical Quarterly*, 55 (220): 385-402

Oberman, Keiron. 2016. 'Immigration as a Human Right' in Sarah Fine and Leah Ypi eds. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press

O'Keefe, M.,K. Klebe, A. Stucker, K. Sturm, and W. Leggett, 2010. *One Year Longitudinal Study of the Psychological Effects of Administrative Segregation*, Washington, USA: US Department of Justice

Otsuka, M. 2003. *Libertarianism without Inequality* Oxford: Oxford University Press

Otsuka, M. 2009, 'Moral Luck: Optional, Not Brute', *Philosophical Perspectives*, 23: 373–388

Pasternak, A. 2011. 'Sharing the Costs of Political Injustices', *Politics, Philosophy, and Economics*, 10 (2): 188-210

Primoratz, I. 1989. 'Punishment as Language', *Philosophy*, 64: 187-205

Prison Reform Trust. 2013. *Prison: the Facts*. London: Prison Reform Trust

Quinn, W. 1985. 'The Right To Threaten And The Right To Punish', *Philosophy & Public Affairs*, 14(4): 327-373

Quong, J. 2009. 'Killing in Self-Defence', *Ethics*, 119(3): 507-537

Sangiovanni, A. 2007. 'Global Justice, Reciprocity, and the State', *Philosophy & Public Affairs*, 35(1): 4-39

Sayre-McCord, G. 2001. 'Criminal Justice and Legal Reparations as an Alternative to Punishment', *Philosophical Issues*, 11: 502–529

Scanlon, T. M. 1986. 'The Significance of Choice', [online] *The Tanner Lectures on Human Values*, Available at: <http://tannerlectures.utah.edu/lectures/atoz.html>)

Scanlon, T. M. 2000. *What We Owe To Each Other*. Cambridge, Mass.: The Belknap Press of Harvard University Press

Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge Mass.: Harvard University Press

Schemmel, C. 2011. 'Why Relational Egalitarians Should Care About Distributions', *Social Theory and Practice*, 37(3): 365-390

Sen, A. 2004. 'Elements of a theory of human rights', *Philosophy & Public Affairs*, 32(4): 315-356

Sen, A. 2006. 'What Do We Want from a Theory of Justice', *Journal of Philosophy*, 103 (5): 215–238

Sher, George. 2013. 'Wrongdoing and Relationships: The Problem of the Stranger' in D. Justin Coates and Neal A. Tognazzini eds. *Blame: Its Nature and Norms*. Oxford: Oxford University Press

Shin, P. 2009. 'The Substantive Principle of Equal Treatment', *Legal Theory*, 15 (2): 149-172

Shoemaker, David. 2013. 'Blame and Punishment' in D. Justin Coates and Neal A. Tognazzini eds. *Blame: Its Nature and Norms*. Oxford: Oxford University Press

Simmons, A. J. 2010. 'Ideal and Nonideal Theory', *Philosophy & Public Affairs*, 38(1): 5–36

Smart, J.J.C. 1961. 'Free Will, Praise, and Blame', *Mind*, 70: 291–306

Smart, John J. C. 1973, 'An Outline of a System of Utilitarian Ethics' in J. Smart & B. Williams, *Utilitarianism: For and Against*. Cambridge: Cambridge University Press

Smith, A. 2007. 'On Being Responsible and Holding Responsible', *The Journal of Ethics*, 11: 465–484

Smith, Angela. 2013. 'Moral Blame and Moral Protest' in D. Justin Coates and Neal A. Tognazzini eds. *Blame: Its Nature and Norms*. Oxford: Oxford University Press

Song, Sarah. 2016. 'The Significance of Territorial Presence and the Rights of Immigrants' in Sarah Fine and Leah Ypi eds. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press

Southwood, Nicholas and Geoffrey Brennan. 2007. 'Feasibility in action and attitude' in J. Josefsson D. Egonsson eds. [online] *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*, Available at:
<http://www.lunduniversity.lu.se/lup/publication/1f14bc1d-bb9d-450b-a5cf-5b9ade6044e2>

Stanley, Emma and Stuart Byrne. 2000. 'Mothers In Prison: Coping With Separation From Children' Paper presented at the *Women in Corrections: Staff and Clients Conference* convened by the Australian Institute of Criminology in conjunction with the Department for Correctional Services SA. Also online, Accessed at:
http://aic.gov.au/media_library/conferences/womencorrections/stanbyrn.pdf

Stemplowska, Z. 2008. 'What's Ideal about Ideal Theory?', *Social Theory and Practice*, 34 (3): 319–40

- Stemplowska, Z. 2016. 'Doing more than one's fair share', *Critical Review of International Social and Political Philosophy*, 19 (5): 1-18
- Stiltz, A. 2009. *Liberal Loyalty: Freedom, Obligation, and the State*. Princeton: Princeton University Press
- Strawson, Peter. 1982. 'Freedom and Resentment' in Gary Watson ed. *Free Will*. Oxford: Oxford University Press
- Sher, G. 2006. *In Praise of Blame*. Oxford: Oxford University Press
- Tadros, V. 2009. 'Poverty and criminal responsibility', *The Journal of Value Inquiry*, 43(3): 391-413
- Tadros, V. 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press
- Tadros, V. 2012. 'Replies', *Jerusalem Review of Legal Studies*, 5(1): 89–109
- Tadros, V. 2013. 'Responses', *Law and Philosophy*, 32(2-3): 241-325
- Tasioulas, J. 2006. 'Punishment and Repentance', *Philosophy*, 81: 279-322
- Tan, K. 2004. *Justice Without Borders: Cosmopolitanism, Nationalism and Patriotism*. Cambridge: Cambridge University Press
- Thomson, J. J. 1991. 'Self-Defence', *Philosophy & Public Affairs*, 20 (4): 283-310
- Thomson, J. J. 1999. 'Physician Assisted Suicide: Two Moral Arguments', *Ethics*, 109(3): 497-518
- Valentini, L. 2012. 'Ideal vs. Non-ideal Theory: A Conceptual Map', *Philosophy Compass*, 7 (9): 654-664

Walmsley, R. 2016. *World Prison Population List 11th Edition*. London: Institute for Criminal Policy Research.

Watson, Gary. 1993. 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme' in J. M. Fischer and M. Ravizza eds.. *Perspectives on moral responsibility*. Ithaca, NY: Cornell University Press

Wellman, C. H. 2012. 'The Rights Forfeiture Theory of Punishment', *Ethics*, 122(2): 371-393

Wellman, Christopher. 2016. 'Freedom of Movement and the Rights to Enter and Exit' in Sarah Fine and Leah Ypi eds. *Migration in Political Theory: The Ethics of Movement and Membership*. Oxford: Oxford University Press

Wellman, C. and Philip Cole. 2015. *Debating the Ethics of Immigration: Is There a Right to Exclude?* Oxford: Oxford University Press

Ypi, Lea. 2013. 'Cosmopolitanism Without If and Without But' in Gillian Brock ed. *Cosmopolitanism versus Non-Cosmopolitanism*. Oxford: Oxford University Press