
POLITICAL THEORY AND THE CRIMINAL LAW

MATT MATRAVERS*

1 INTRODUCTION

This chapter asks what contractarian political theory can tell us about the place of the criminal law, and of criminals, in a liberal society. It begins by distinguishing two forms of contractarian thinking: one owed to the mutual advantage tradition of Hobbes; the other to the impartialist tradition of Kant. The argument is that both can underwrite a system of criminal law, and of punishment, for similar reasons.

However, at this point the traditions diverge. The mutual advantage tradition allows—indeed, commends—the transmission of natural inequalities into just outcomes. The Kantian tradition, in particular in Rawls, does not. Instead, it posits a position of fundamental equality from which, in distributive justice at least, we move only when, in Rawls’s language, we agree ‘to share one another’s fate’. Much

* Earlier versions of this chapter were given to the UK ALPP Conference at York and to a gathering of lawyers and philosophers at Rutgers. I am grateful to Matt Kramer, Antony Duff, and Stuart Green for their invitations to try out some of these ideas, and to the audiences on both occasions for constructive comments. I am also grateful to Antony and Stuart for written comments. I am sure I have not met all their objections, but I have made so many changes (and, I hope, improvements) in response to their comments that I have not noted each one. The first inklings of the ideas discussed in the paper, were batted back and fore with my colleague, Sue Mendus, to whom (as ever) I am indebted.

of the argument then concerns whether it is possible to (and whether there is reason to) apply this Rawlsian insight to the retributive sphere and what it would mean to do so.

The argument proceeds at a high level of both abstraction and generality. To some extent this is inevitable, it seems to me, given that political theory tends to operate at some distance from the immediate and practical and given that it is probably right to do so. Political theory tends to generate general, abstract principles that may well not be fine grained enough to answer the question of whether we should do *X* or *Y* when confronted with that choice of policy options.¹ Nevertheless, the paper concludes with some reflections on the attitude citizens should take to one another and to the criminal justice system that they reflectively endorse in the contract.

2 THE CRIMINAL LAW IN A LIBERAL STATE

2.1 Methodological preamble

One method of enquiry into the philosophical foundations of the criminal law, or (only slightly) more modestly into the place and nature of the criminal law in a liberal state, is to ask what people would agree to in some hypothetical choosing situation (that is, to invoke contractarianism). Contract theory is often alleged to hide substantive normative commitments in itself—and has variously been criticized as sexist, speciesist, and normatively individualistic—but contract theory, suitably understood, need not be any of these things; it all depends on the specification of the choosing situation and of those who do the choosing.²

I have argued elsewhere that rational, self-interested choosers could endorse rules of cooperation as moral norms, but in order so to do they would (given certain

¹ This is not to say that particular political theorists have not contributed to specific policy debates; clearly they have. However, to use the techniques of critical rational enquiry to examine questions of what we should do differs from deriving what we should do from some abstract political theory. Brian Barry is a useful case in point. His *Culture and Equality* and *Why Social Justice Matters* both offer specific policy proposals. These are said to ‘satisfy the reasonable rejectability test’ as developed in his more theoretical *Justice as Impartiality*. This claim can be disputed, but even if it is true, it does not mean that these are the only policies that could satisfy the contractarian test (and Barry does not claim that they are). See BM Barry, *Justice as Impartiality: Volume 2 of A Treatise on Social Justice* (Oxford: Clarendon Press, 1995); *Why Social Justice Matters* (Cambridge: Polity Press, 2005); *Culture and Equality: An Egalitarian Critique of Multiculturalism*, (Cambridge: Polity Press, 2001).

² A particularly useful account of constructivist theorising in general is in BM Barry, *Theories of Justice: Volume 1 of A Treatise on Social Justice* (Hemel Hempstead, Herts: Harvester-Wheatsheaf, 1989) 262–84.

plausible background factors obtaining) also have to endorse a system of punishment.³ The function of the system of punishment would be to address an assurance problem (parties can only be expected to endorse constraints on their pursuit of self-interest if they can be assured that others will do so too); to communicate to the offender the significance of the commitment to morality (both individual and general); and, where necessary, to educate the offender so as to enhance his ability to engage with others on moral terms. Rationality does not compel persons to endorse the rules of cooperation as moral, it cannot compel acceptance of the system of punishment, but it does not undermine this rationale for these practices.

Clearly, the particular form of contractarianism I defend contains substantive normative and methodological commitments, and these are of course not uncontroversial.⁴ However, for present purposes, I intend to try to work with general, (I hope) plausible, and less controversial claims in order to develop an account of some aspects of the criminal law that could be accepted at least by most liberals. To give just one example of the kind of consensus with which the paper tries to work, consider the (some would say special) need to justify social arrangements to those individuals who do worst under those arrangements.

According to the contractarian position I defend, it makes no sense to think that justification is owed to others. Rather, the (primary) problem of morality is how to justify to oneself moral constraints on the pursuit of one's self-interest. According to the much more common liberal contractarianism of, for example, Rawls, Barry, and Scanlon, the issue is what we owe and can justify to each other.⁵ This dissimilarity reflects a fundamental difference between my argument and that of liberal egalitarians about the nature and scope of morality. Nevertheless, on both accounts it is plausible to think that there is something special about those who do worst under any given set of social arrangements. For on both accounts there is some set of social arrangements, which could be otherwise, under which some persons do worse than others. It seems natural to think, from the one perspective, that some justification needs to be offered to these people in particular,⁶ and, from the other, that the problem of justifying to oneself a commitment to constrain the pursuit of one's self-interest in accordance with norms that are part of a set of social arrangements

³ See M Matravers, *Justice and Punishment: The Rationale of Coercion* (Oxford: Oxford University Press, 2000).

⁴ Some of the more controversial commitments are usefully brought out, and subject to criticism, in R Shafer-Landau, 'Matt Matravers, Justice and Punishment: The Rationale of Coercion' (2004) 114 *Ethics* 361.

⁵ David Gauthier usefully compares the ideas of justification to self and to others in his 'Political Contractarianism' (1997) 5 *Journal of Political Philosophy* 132.

⁶ Of course, Rawls goes further than this and gives the worst-off group a special status and the equivalent of a veto on the agreed set of social arrangements, but whether he is justified in doing so is a moot point. Whether he is or not, throughout *A Theory of Justice* it is clear that Rawls thinks that there is a special need to justify inequalities to those who do worse out of them. See JB Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).

under which one does worse than others, will be particularly acute. As the argument proceeds, more will have to be said about these two forms of contractarianism, not least when they recommend different determinate answers to pressing legal and political problems.

The reference (in the previous paragraph) to the fact that social arrangements could be otherwise hides one substantive commitment. This is to the claim that contractarianism builds on natural facts about us and the world, but that these facts are normatively ‘inert’. That is, natural facts themselves are (in Rawls’s words) ‘neither just nor unjust’, and they do not dictate the shape of agreed social arrangements. Rather, we must recognize that ‘the social system is not an unchangeable order beyond human control but a pattern of human action’.⁷ That system can be just or unjust and it is up to us which it is. For example, the fact that human beings normally feel pain is not in itself just or unjust, right or wrong. It is just a fact about the psycho-physical make up of human beings. But a set of institutions that needlessly allows the inflicting of pain on some set of human beings may well be unjust and, insofar as it is, and we reflectively endorse it in a contractarian thought experiment, we are complicit in that injustice.

I take it that this constructivist commitment need not be shared by all contractarians. For example, a theorist could invoke contractual thinking merely as a heuristic device and use the contractors as means to reflect on the innate moral order of the universe as given by God. Thus in stipulating it as a commitment I am ruling out certain forms of natural law theory (and certain understandings of such things as ‘wrongs in themselves’). However, I do not think this is particularly significant. There are other forms of natural law theory that could fit the proposed contractual scheme, but even if that were not the case, the constructivist commitment is something that is shared by those whose liberal dispositions are the target of this paper.

It is worth making one final remark about the use of contract theory in this paper. It might be thought that if the argument is going to rely on either shared, uncontroversial, premises, or premises that are explicitly spelt out so as to make differences between contractual accounts transparent, then it could proceed without any reference to ‘contract’. The argument would simply move from premise to conclusion. What then is the purpose of the language of contract? Of course, this is a common question asked of contract theory and a line of criticism with impeccable philosophical credentials has it that talk of contracts is (at best) little more than unnecessary window dressing.⁸ The answer, the best account of which is given by Samuel Freeman, is that contract theory describes a form of rational reflection. Anyone, at any time, can ‘enter’ the contract simply by thinking in the manner described in the construction of the hypothetical choosing situation. That person then reasons

⁷ All quotations from Rawls, *ibid* 102.

⁸ D Hume, ‘Of the Original Contract’ in D Hume, *Essays—Moral, Political and Literary* (EF Miller ed; Indianapolis: Liberty Classics, 1987).

from premise to conclusion; in the rhetoric, she ‘chooses’ principles of justice (or whatever). The notion of contract captures two ideas: first, that if the choosing situation is properly constructed then every person entering the contract and thinking in this way ought to be able to acknowledge the conclusions as both right and ‘theirs’. Second, each agrees only on the understanding that others do, too. As Freeman puts it, ‘the *mutual acknowledgement* of principles... warrants the term “agreement”, and the *mutual precommitment* involved might just as well be called a “contract”’.⁹

2.2 Contractarianism and the criminal law

As noted above, I have argued elsewhere that we can—taking up the standpoint of self-interested agents engaged with one another for mutual advantage—reflectively endorse a system of criminal prohibitions, enforcement and punishment. This is true, too, for Kantian-inspired contractarianism. The reason is that when we ‘contract’ to live together on moral terms—or terms that bind us together as citizens in a well-ordered society—we do so on the basis that others do so, too, and that we can be reasonably assured of their compliance with the terms of agreement. The state’s enforcing of the law—in relation to, for example, taxation—allows each person reflectively to endorse her commitment to the scheme of social cooperation.

As stated, it may not be immediately clear why the assurance problem requires *criminalization* rather than just *regulation*. The short version of the answer is that it does not. However, if the arrangements endorsed by the contracting parties are to be stable and are to avoid some of the well-documented failures of mutual advantage theory, then the parties must affirm the principles by which they agree to be governed as having the imperatival force of moral principles. It is this that makes it a theory of morality rather than of mere cooperation.¹⁰

⁹ SFreeman, ‘Introduction: John Rawls: an overview’ in SFreeman (ed), *The Cambridge Companion to Rawls* (Cambridge: Cambridge University Press, 2003) 1, 19, original emphasis. These ideas—that by reflective endorsement we commit to social arrangements that we endorse as ‘ours’—will be important in what follows.

¹⁰ I believe the above description is sufficiently broad to encompass theories in both the mutual advantage tradition (which includes Hobbes, Gauthier, and the account I defend above) and the Kantian tradition (of Kant, Rawls, Barry, and Scanlon). Differences between the traditions emerge as soon as one digs deeper, in this case, into the nature of the endorsement of the principles as moral. In *A Theory of Justice*, justice and goodness are said to be congruent, so one has reason to endorse the primacy of reasons of justice (relative to prudential reasons of short-term advantage) because to do so is one’s good as a free, rational being. This is the account with which Rawls became disillusioned, and in *Political Liberalism* the parties endorse the principles from within their own comprehensive views in an ‘overlapping consensus’. In Barry, the contractors are motivated by a sense of justice, which amounts to a recognition of the special place that should be accorded to reasons of justice in practical deliberation. The mutual advantage tradition finds it harder to ground the imperatival force of agreed principles given that the contract is built around the idea of advancing individual contractors’ interests. For Gauthier, what makes his theory an account of *morals* by agreement and not *rules*

Failure by one person to adhere to the rules, then, is a moral failure in relation to all others. Moreover, on this account, it is a particularly damaging failure since the scheme works only to the degree that all contractors are willing to make the necessary commitment, which itself depends on their being assured that others will do so too. The appropriate response to such a failure, then, is one of condemnation and not simply of correction. In condemning the offender, the state reasserts the moral value of the agreed principles and reminds the offender of his agreement to abide by those principles. At least, that holds for those who agree to contract by reflectively endorsing the principles as principles that rightfully govern their pursuit of their own advantage. The need to address the assurance problem, and recognition of the prudential reason of agents, underpins an account of hard treatment. The recognition of the moral commitment of the parties to the construction of the community underpins the account of censure and condemnation.

This contractual account of the criminal law is clearly sketchy. The degree to which the procedure can generate precise answers as to the content and scope of the criminal law is an interesting question, but one which I am not going to pursue here. Obviously, given the types of beings that we (human beings) are, the content of the criminal law will concern wrongful harms of certain kinds, and given the kinds of community in which we currently exist, it will also cover certain economic and social spheres. The question I want to pursue concerns the justification of the system as a whole, its relationship to ideas of desert, and the attitude participants should have to one another (including to those who are punished).

3 CONTRACTARIANISM AND NATURAL (DIS)ADVANTAGES

Consider the position of those who are born deaf. Presume (plausibly) that modern societies make deafness more of a disadvantage than it would be under other more simple social arrangements. On the account of contractarianism based on mutual advantage, the critical question for the deaf and the hearing is whether mutually advantageous cooperation is possible and, if so, under what conditions.

by agreement is the idea that it can be rational to adopt a disposition (constrained maximization) such that one is disposed to keep agreements even where immediate self-interest might better be served by free-riding. On my account, rationality cannot quite deliver that, so what is required is an 'existential leap' on the part of the contractors see Barry, *Justice as Impartiality*; D Gauthier, *Morals by Agreement* (Oxford: Oxford University Press, 1986); Matravers, *Justice and Punishment*; Rawls, *A Theory of Justice*; J Rawls, *Political Liberalism: Expanded Edition* (New York: Columbia University Press, 2005).

For the mutual advantage theorist, each contractor must advance his interests through cooperation relative to the baseline of non-cooperation. What this may mean is that no mutually advantageous ‘deal’ can be done, in which case those left outside cooperation are left ‘beyond the pale’—the protection of—morality.¹¹ Where a deal can be done, the bargaining solution will reflect the unequal starting points of the parties and, in this sense, natural advantage and disadvantage will be transmitted through the bargain into the outcome. For many, this is one reason why the mutual advantage tradition in contractarianism is flawed (and worse¹²). For its critics, the job of justice is (at least in part) to protect the weak not least by negating the effects of natural inequalities. This is, of course, at the heart of Rawlsian contractarianism.

As already noted, for Rawls the fact that some people are born, for example, deaf is ‘neither just nor unjust’. What can be just or unjust is the social system given the fact that some people are hearing and others deaf. For Rawls, a just social system will not reflect, but will neutralize, natural advantages and disadvantages. ‘In justice as fairness’ as Rawls memorably puts it we ‘agree to share one another’s fate’.¹³ That is to say, for Rawls the question of whether the hearing would advance their interests by excluding the deaf from cooperation is not relevant. In reflecting on principles of justice, we realize the morally arbitrary nature of natural (dis)advantages and commit to live together in a well-ordered society. We initially share one another’s fates—in this example, we share the fate of the deaf—by excluding knowledge of all personal information in the choosing situation (that is, by imposing the veil of ignorance on the parties in the choosing situation). By doing this we recognize the morally arbitrary nature of natural starting points and ensure that those (unequal) starting points are not automatically transmitted into unequal outcomes (even in the case where allowing natural inequalities to be reflected in the outcome of the contract would be to the advantage of some of us).¹⁴

Rawls’s account of the moral arbitrariness of both social and natural (dis)advantages underpins his radical account of equality of opportunity. If justice requires equality of opportunity, then we should ignore natural (dis)advantage just as we

¹¹ Gauthier (in)famously describes ‘animals, the unborn, the congenitally handicapped and defective’ as beyond the pale of a morality tied to mutual advantage. See Gauthier, *ibid* 216.

¹² Will Kymlicka, for example, describes the mutual advantage tradition as not supplying an alternative theory of morality, but ‘an alternative morality’, and Brian Barry characterizes the position as ‘morally pathological’. See Barry, *Justice as Impartiality* and W Kymlicka, ‘The Social Contract Tradition’ in P Singer (ed), *A Companion to Ethics* (Oxford: Blackwell, 1991).

¹³ Rawls, *A Theory of Justice*, 102.

¹⁴ Those who are suspicious of Rawls’s (methodological) individualism sometimes balk at the centrality he accords to ideas such as ‘fraternity’ and to his invoking (here) of the idea of ‘sharing’ one another’s fate. However, it needs to be remembered that the language of self-interested rational choice is relevant in Rawls only once the original position (the choosing situation) is defined. The characterization of the original position—including the thick veil of ignorance—reflects deep moral convictions (particularly a commitment to fundamental equality).

ignore social (dis)advantage. In general terms, this gives rise to Rawls's non-desert based principles of distributive justice (in which legitimate expectations replace desert), which ensure equality in the distribution of basic rights, equality of opportunity, and inequalities in the distribution of social and economic goods only insofar as those inequalities maximally benefit the least well-off.¹⁵

What this means for deaf and naturally disadvantaged people is something like this: they are, of course, included in the contract and thus are entitled to the protections of the first principle (that is, to equal basic liberties). They are also entitled to equality of opportunity when it comes to the chance to enjoy social and economic inequalities. Quite what this means in practice will be complicated, but the idea is clearly that opportunities should be made available to all wherever it is reasonable to do so. Despite these protections, it may well be that some of those who are naturally (and/or socially) disadvantaged still end up in the economically worst-off group. However, if that is the case, they do not do so because they are less able or less deserving. Rather, there are inequalities in the system—inequalities that allow others to do better than them—only because those inequalities maximally benefit the position of the worst-off. In this sense, too, we share one another's fates in that natural features of persons play a role in distribution that is constrained by the system as a whole. If the talented do well it is only because by doing well they benefit the worst-off.

What is the relevance of this to questions about the criminal law? On one account, none, because retributive justice (broadly conceived) is different from distributive justice and the arguments in the one sphere do not translate into the other.¹⁶ However, I have argued elsewhere that this is not the case: the same concerns should underpin our analysis of the basic structure whether in relation to retributive or to distributive justice.¹⁷ I will not rehearse that argument here, but I hope that the discussion of the examples that follow will make the case seem plausible (even if not proven).

Allowing that Rawls's broad approach can be applied in the retributive sphere means departing from Rawls's own assumptions of ideal theory and full compliance. It also means applying his account to a question to which he thought it did not apply. That is, one cannot be true to the text and ask what the persons in the original position would choose in relation to criminal justice because Rawls did not believe the question to be one appropriately dealt with in this way.¹⁸ We can depart from Rawls whilst borrowing from his account, though, which is the project here. This

¹⁵ See particularly Rawls, *A Theory of Justice*, §48.

¹⁶ This is the argument of Rawls himself and Samuel Scheffler: S Scheffler, *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought* (Oxford: Oxford University Press, 2001).

¹⁷ See M Matravers, 'Mad, Bad, or Faulty: Desert in Distributive and Retributive Justice' in C Knight and Z Stemplowska (eds), *Responsibility and Distributive Justice* (Oxford: Oxford University Press, forthcoming).

¹⁸ Thus, to ask what a Rawlsian theory of punishment would be like, and to try to answer that question by trying to apply Rawls's theory directly, strikes me as (at best) an invitation to perform intellectual contortions of a quite demanding kind and (at worst) a straightforward mistake. That is not to

means avoiding the details, but hanging on to the moral commitments that drive the theory.

The place to start, then, is with the construction of the original position. As with natural inequalities and all other personal factors, one's disposition to criminal behaviour would be included in the veil of ignorance (that is, it would not be known to the contracting parties). I have argued above that the contracting parties would have reason to choose—reflectively to endorse—a system of criminal law. What the veil of ignorance adds to that is that one will not know one's risk of falling foul of that law, either because of a disposition to criminal behaviour or because of circumstance, or for that matter mistake (on your part or on that of the system). As a risk-averse contractor, then, one has reason to choose a system in relation to criminal behaviour that is, in the words of the sometime UK Prime Minister Tony Blair, 'tough on crime and tough on the causes of crime'. That is, one will choose a system of criminal law (for the reasons given above), but surround that system with protections—including, but going beyond those of equal legal rights, etc—that reduce the prospect of one being subject to punishment (this will hold even where punishment only affects the offender, but will be even more important given the 'spill over' effects that punishment has on family, friends, job prospects, etc, in the real world).

Clearly this account needs to be unpacked, and there will be those who will have already balked at being asked to consider natural facts such as talent or disability together with a disposition to criminal behaviour, but it is worth pushing on a little further before considering possible criticisms.

Recall, for the Rawlsian liberal egalitarian, natural facts are what I called morally inert, or what Rawls calls morally arbitrary. This means that we begin with a conception of the members of the society as fundamental moral equals and design the basic structure on that basis. Once the basic structure is in place, our ordinary social practices continue, but on the basis of the principles chosen to govern that structure. 'Thus', as Rawls puts it:

it is true that as persons and groups take part in just arrangements, they acquire claims on one another defined by the publicly recognized rules. Having done various things encouraged by the existing arrangements, they now have certain rights, and just distributive shares honor these claims. A just scheme, then, answers to what men are entitled to; it satisfies their legitimate expectations as founded upon social institutions.¹⁹

Put more informally, the argument is this: 'natural' people (so to speak) differ in being more or less talented and in things such as their gender and skin colour. In deciding the principles of distribution of rights, political liberties, and economic and social goods, these things are arbitrary because none is connected to a pre-justicial notion of moral worth. Therefore, they do not figure in the principles of distribution.

say that nothing interesting results. For example, see S Dolovich, 'Legitimate Punishment in Liberal Democracy' (2003–04) 7 *Buffalo Criminal Law Review* 307.

¹⁹ Rawls, *A Theory of Justice*, 311.

However, once those principles are in place, the principles may (for reasons of, for example, efficiency) encourage the talented into well remunerated professions such as brain surgery (one does not want a talentless brain surgeon). In that case, the trainee brain surgeon will develop an expectation that her training, if successful, will be rewarded and that expectation is legitimate.²⁰

To apply this to the criminal law, then, would be to insist that inert natural facts ought not to dictate the shape of the principles of (retributive) justice. However, once these principles are in place, such facts may well play a role in where, within the scheme, a given person ends up. Just as the talented (and socially lucky) will tend towards the better-off groups—not because justice requires rewarding the talented, but because rewarding the talented maximally benefits the least well-off—those who are disposed to break the criminal law (for whatever reason) will tend towards the group who are punished, but again not because justice requires principles that punish those who act on such a disposition, but because only by punishing them will the system provide the assurance needed to be stable.²¹

Extending Rawls's argument in this way is something for which one can find some encouragement in the text. After all, Rawls famously denies the connection between justice and moral desert. He writes, in a passage that follows directly from that quoted above, 'but what they [persons taking part in just arrangements] are entitled to is not proportional to nor dependent upon their intrinsic worth. The principles of justice that regulate the basic structure and specify the duties and obligations of individuals do not mention moral desert, and there is no tendency for distributive shares to correspond to it.'

It seems to me that the argument offered above is plausible and is, indeed, one that Rawls should have left open. However, as already noted, and as the reference to '*distributive* shares' in the last quotation makes clear, Rawls is explicitly committed to the claim that distributive justice is not 'somehow the opposite of retributive justice'.²² By this, Rawls seems to mean that principles of retributive justice can legitimately refer to the pre-justicial moral worth of the person (or the person's actions), and thus these features of the person are not morally arbitrary.²³ This is a common

²⁰ The legitimacy of this claim is disputed by some including GA Cohen (see GA Cohen, 'Incentives, Inequality, and Community' in S Darwall (ed), *Equal Freedom: Selected Tanner Lectures on Human Values* (Ann Arbor: The University of Michigan Press, 1995) 331–97).

²¹ The account offered in the last few paragraphs owes a great deal to a discussion with Jo Wolff at the UK ALPP Conference. What is remarkable is the degree to which the position recalls Rawls's justly famous defence of rule utilitarianism in 'Two Concepts of Rules'. In short, the overall purpose of the system of punishment is given (primarily) by the need for assurance. Once established, the rules governing the application of punishment are retributive. See J Rawls, 'Two Concepts of Rules' in JB Rawls and S Freeman (eds), *Collected Papers* (Cambridge, MA: Harvard University Press, 1995 (1999)).

²² Rawls, *A Theory of Justice*, 314.

²³ It should be said that Rawls writes of retributive justice in different ways in different places. Sometimes, he comes close to the post-justicial legitimate expectations view described here. Consider the following passage (235): 'A Legal system is a coercive order of public rules addressed to rational persons for the purpose of regulating their conduct and providing the framework for social

response and, if the argument above is to have any plausibility, something needs to be said about the (dis)analogy I am drawing between distributive and retributive justice (and so between, for example, the case of the talentless person who ends up in the worst-off group and the person disposed to break the criminal law who ends up being punished).

I have argued so far that we might think of ‘sharing each other’s fate’ in the retributive sphere in a way that is analogous to the way in which we do that in the distributive sphere. Retributive questions are asked in the original position, and the people in that position do not know their tendency to disobey, or the likelihood that they will fall foul of, the criminal law. They have reason to choose a system of criminal law, but they also have reason to surround that system with protections that will reduce the likelihood of their being punished. Perhaps more radical than that is the suggestion that just as those who do well or badly in distributive terms can only be properly thought of in terms of legitimate entitlement and not desert—one does well because by doing well one maximally benefits the least well-off and one does badly only because positions of relative disadvantage exist only so that the least well-off can be as well-off as possible—so those who are punished are punished not because they (pre-justicially) deserve it but because they are entitled to it under a just scheme in which punishment has some other, non-desert based, rationale.

For some—as we have seen, including Rawls—this position is unsustainable. One version of the objection can be captured if one thinks of one way in which the position of the talentless person who ends up in the worst-off group and the person disposed to criminal behaviour who ends up punished seem disanalogous. In addressing the talentless, it is not just Rawls who might hold that there is something objectionable in saying, just like that, ‘the explanation and justification for your being worse-off than others is that you are talentless’. However, in the case of the punished, it seems enough to say ‘the explanation and justification for your being punished is that you broke the law’. The difference seems to be one of responsibility. According to a well-established liberal position, it is not justifiable to hold people to account, and to make them pick up the burden, for things over which they had no control. The talentless did not choose to be talentless, but the criminal, *ex hypothesi*, did choose to break the law.

According to one reading of Rawls, ‘moral arbitrariness’ depends on not being responsible. So, factors such as one’s height, intelligence, and talents are morally arbitrary because they are unchosen. On this reading, it is thus because they are unchosen that these features of people are hidden from the view of the people in the original position by the veil of ignorance. If so, and if criminal behaviour is chosen, then one’s tendency to criminality would have to be known to the people in the

co-operation. When these rules are just they establish a basis for legitimate expectations. They constitute grounds upon which persons can rely on one another and rightly object when their expectations are not fulfilled.’

original position and thus would be excluded from the realm in which we share one another's fates.

Although there is some textual evidence for this view, it is not Rawls's.²⁴ Rather, for Rawls the veil of ignorance captures the commitment to the idea that persons are fundamentally equal. Moral desert as a basis for justice is rejected because there is no sensible way of moving from desert to distributive outcomes (an argument that seems to me to be at least as plausible in the case of retributive judgements as distributive ones).

Nevertheless, the responsibility sensitive position is so widespread that it is worth considering its application here before, finally, considering how and in what ways we share the fate of others in the retributive sphere.

3.1 Responsibility and natural (dis)advantage

Consider someone who is disabled, but not visibly so, who enters a two-storey building and asks that the janitor come out to activate the elevator. The janitor may well ask why he should be inconvenienced, but on being told of the disability he would presumably accept that the person has good reasons for needing the elevator and would act accordingly. Contrast this with an agent predisposed to aggression. The case here is more complicated. Assume the agent to have assaulted someone as a result of a perceived (or real) minor slight. In this case, we hold the agent responsible. If the agent explains that he is genuinely incapable of acting otherwise—he has a disorder such that he loses control over himself completely—then that judgement is revised. However, if his explanation is that he is simply the kind of person who is quickly angered and acts on that anger, then we do not think that an excuse. The agent is the subject of our reactive attitudes and is held responsible for his aggressive act.

According to mainstream compatibilist accounts, in both cases the principal actors act for reasons, but the disabled person has good reasons, and reasons that underwrite his not fully bearing the costs of his disability. The aggressive individual also acts on reasons, but it on the basis of so-doing that he is rightly held to account and asked to pick up the bill for his actions.

Now, one strategy in response to this might be to deny that compatibilism can do the work asked of it. One might here appeal to the (mis)reading of Rawls that has it that all features of individuals are unchosen and so undeserved up to and including the reasons on which we act. If so, mainstream compatibilism does not adequately justify our reactive attitudes and the practices of blaming and punishing (as well as praising and rewarding) that are associated with them. Compatibilism, arguably,

²⁴ See my M Matravers, *Responsibility and Justice* (Cambridge: Polity Press, 2007).

shows that some form of responsibility is compatible with the truth of ‘the causal thesis’, but it is a hollow form of responsibility when what we seek is something much more robust to underwrite those practices.²⁵ Although interesting, I want to put this response to one side.

The second response admits the relevance of responsibility, but only after the system of justice is established. This is, of course, simply to return to the entitlement system commended by Rawls in relation to distributive justice. After all, the student who works hard to be a brain surgeon acts responsibly and is rightly praised. He is not, though, rewarded directly because of his talents (including the talent of working hard), but because the system that rewards people like him maximally benefits the worst-off. The reason for this, as we have seen, is not because of some incompatibilist premise that Rawls failed to make explicit. It is that there is no justification for allowing inequalities in natural facts (or social luck) to be reflected in the principles of justice. The initial position of equality is fundamental. Stepping away from equality can be justified in some cases—not in matters of basic rights, but in the economic realm—but on the grounds of advancing the position of the worst-off and not on grounds of moral desert. Even were we to attempt to find some proxy for moral desert such as the willingness to engage in conscientious effort, we would be defeated since even that, Rawls notes, is as much to do with one’s upbringing as one’s natural talents.²⁶

4 ‘SHARING ONE ANOTHER’S FATE’ IN RETRIBUTIVE JUSTICE

To re-cap: I have argued that if we extend Rawls’s arguments about the irrelevance of natural starting points and the social lottery from distributive to retributive justice two important things result. First, and unlike the mutual advantage tradition, we exclude the possibility of placing some—the congenitally dangerous, say—outside the protection of justice. Since, to do so might well be mutually advantageous for those who remain, I take this to be an initial aspect of what it is to share one another’s

²⁵ See, for example, G Strawson, ‘The Impossibility of Moral Responsibility’ (1994) 75 *Philosophical Studies* 5; and for a discussion Matravers *Responsibility and Justice*, *ibid.* The phrase ‘the causal thesis’ is taken from (but possibly not original to) Scanlon to capture the claim that ‘all our actions have antecedent causes to which they are linked by causal laws of the kind that govern other events in the universe’. TM Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998) 250.

²⁶ Rawls, *A Theory of Justice*, 102.

fate; it is to include all—whatever their ability (or inability) to contribute to the social product. Second, the justification for the system of punishment as a whole will lie primarily in its providing the necessary assurance to make contracting (reflective endorsement) possible. Once in place, the system of criminal law and punishment will give rise to legitimate expectations that the system must honour. The arrangements that give rise to these expectations will be responsibility-sensitive (since it is only by punishing only the guilty that the purposes of the system can be achieved) such that those who are punished can be told that they are being punished because of their actions against a background in which a great deal is done to mitigate natural, and eliminate social, causes of crime. However, in reply to the question of why there should be a system of punishment at all, the answer does not appeal to moral desert—to the idea that the system exists to give pre-justicially deserving people the suffering they deserve—but to the overall good achieved by the system including the good of the person being punished.

I take it that this is the analogue of the scheme of distributive justice in which all (citizens) are included and in which unequal positions only exist to achieve an overall outcome (the best position for the worst-off group). Thus, in response to the question of why a given person should be less well-off than someone in another position, an explanation can be given in terms of that person's talents, abilities, hard work, and their resulting ability to do 'the various things encouraged by the existing arrangements'. However, asked why unequal positions exist at all (which allow rewards for those things that are encouraged), the answer does not appeal to moral desert, but to the overall position of the least well-off. In this sense, too, we share one another's fate.

However, the idea of 'sharing one another's fate' also seems to have a personal dimension that is missing from the above analysis. What, it might be asked, is it to view one another in this light in the retributive domain? To focus on this it might be worth posing a challenge. Recall the disabled person asking for a lift at some inconvenience to the janitor. Given his disability, it is reasonable to ask for the lift and (given certain background facts) it is reasonable that the 'cost' of the lift is somehow shared. More generally, we share the fate of the disabled by paying into general taxation some of which is then spent in providing them with support, ensuring equal opportunities, where reasonable, etc. Compare this, then, with someone predisposed to aggression who enters the building and asks the janitor to take a beating so that he can relieve his aggressive tendencies. Clearly, here, we do not think that the janitor (or anyone else) should comply. Sharing one another's fate must not be reduced to this.

The question of general taxation, though, is more difficult. Obviously, we do not think that we should pay into general taxation to support lifestyles that are criminal or aggressive. However, we do think that it is a reasonable use of general taxation to support those who have fallen foul of the criminal law to rebuild their lives.

Moreover, as I have already argued, we think it is a reasonable use of general taxation to reduce the social causes of crime, and to provide early intervention for those who seem to be set on a path of criminality.

These things, I think, speak to a way in which we share the fate of others at a personal level. Underpinning the position is a sense of the contingency of natural starting points, upbringing, and social arrangements. And this is the final sense of sharing one another's fate that I wish to consider.

It is, as already noted, an aspect of the constructivist position that social arrangements could be otherwise, and that in reflectively endorsing them we together take responsibility for them. It is, of course, also the case that a given set of social arrangements—even a just set—will suit some more than others. This is in part simply a matter of circumstance. For example, given some technological advances, certain skills will become redundant and those who have cultivated those skills will find their relative position in the market reduced. Other skills may, for contingent reasons, be rare and so their market value enhanced. This does not mean that we should not admire the holder of some rare skill, but it does surely mean that our, and her, attitude to her place in the market should be tempered by an awareness of the role of chance and contingency. We do not have to think of ourselves as entirely the handmaidens of fate to nevertheless appreciate that things could easily be, or have been, different even in a just world.

I take it that the same is true in the retributive sphere. Of course, here choice is more important. The existence of other people with similar skills alters one's market position and that is something that is out of one's control (at least in the first instance). The decision to commit a criminal act is different. And yet, even in a just world, the path that leads to criminality is one that is strewn with contingent features. The modern world provides for many people a confusing, fast-paced, highly stressed environment. It creates criminogenic opportunities and provides temptations in the form of highly valued goods. This is not unjust; it is (as the colloquial saying has it) 'just life'. This does not, on the Rawlsian inspired account offered above, negate responsibility, but it does alter the attitude we have to those who find the temptation to act criminally too strong. If so, then perhaps we should endorse Scanlon's view that whilst there are circumstances in which we can justifiably blame and condemn people, our attitude when we do so 'should not be "You asked for this" but rather "There but for the grace of God go I"'.²⁷ If so, then the final sense of sharing one another's fate is not to think of criminals as different from the rest of us—as 'them' as against 'us'—but instead as just like us only perhaps less fortunate in either their natural (dis)advantages, their social upbringing, or their 'fit' with the world that surrounds them.²⁸

²⁷ Scanlon, *What We Owe to Each Other*, 294.

²⁸ Interestingly, Rawls himself hints at such a possibility when he writes towards the very end of *Theory of Justice* (576) that there may be some for whom 'being disposed to act justly is not a good'. If

5 CONCLUDING REMARKS (OR TWO CAVEATS)

The position for which I have argued in the bulk of this chapter flows, I believe, fairly naturally from its Rawlsian origins. For those who are repulsed by the mutual advantage tradition's transmission of natural inequalities into just outcomes, and/or for those who are impressed by Rawls's deep sensitivity to chance and contingency, the argument should have some attraction. However, for others, the conclusions, far from being compelling, may take the form of a *reductio*. If this is what Rawls leads to, they may say, then so much the worse for Rawls. This chapter does not take a position on this. As noted above, my own commitments are to the mutual advantage tradition, which I believe must incorporate something of the personal dimension of sharing one another's fate described above, but in which the scope of who counts as a relevant 'other' is constrained by their ability to bring something advantageous to the contract.

For those who are convinced—or for whom the argument has at least some purchase—the pressure comes from a different direction. The position requires one to face up to the consequences of chance and contingency—not just in the ways in which they affect us directly, but in that even a just world is one that disadvantages some—whilst maintaining individual responsibility for actions. This strikes me as difficult to do, and the step from 'there but for the grace of God go I' to thinking responsibility irrelevant to justice is a short one.²⁹

there are many such people, Rawls writes, then 'penal devices will play a much larger role in the social system'. But, of these people, he adds, 'one can only say: their nature is their misfortune'. The emphasis on not seeing offenders as 'them' to be contrasted with the law-abiding 'us' is emphasized throughout Duff's writing.

²⁹ It is important to note that this is not an argument about doing (retributive) justice in a (distributively) unjust world. That is, of course, an important topic, but the point here is that even a just world is one that is chosen (or reflectively endorsed) and is one in which some people will fare less well than others. Social structures—even just ones—turn some natural facts into social disadvantages. That is no-one's fault, but it is something that on the Rawlsian account we all share. In this very loose sense, we can be thought (non-culpably) complicit in crime even while we hold the criminal responsible. If so, I think we can learn from the 'doing justice in an unjust world' literature particularly with respect to the attitude we should take to criminal institutions (and criminals). We should always be hesitant about punishment and its justification and when we use it we should always be conscious of a different way the offender and the social world could have been. If this sounds to some readers like bleeding heart liberalism, then Rawlsians should make no apology for that. Even without distributive injustice, this is a realm in which 'we can say that the hearts of bleeding hearted liberals have good reason to bleed' V Tadros, 'Poverty and Criminal Responsibility' (2009) 43 *Journal of Value Inquiry* 391, 413.