

4

Guilty Pleas, Sentence Reductions, and Non-punishment of the Innocent

ZACHARY HOSKINS*

I. Introduction

It is standard practice in the United Kingdom, the United States and other common law countries to reduce criminal sentences in response to guilty pleas. Sentence reductions are intended as incentives to guilty defendants to plead guilty as early in the process as possible (see, eg, Sentencing Council 2017). Guilty pleas are thought to be beneficial in various ways: they may help to reduce victims' suffering, they can save victims and witnesses from having to testify, but the most often cited benefit is that they save time and money that would otherwise be spent on investigations and trials. As US Supreme Court Justice Anthony Kennedy wrote, 'criminal justice today is for the most part a system of pleas, not a system of trials' (*Lafler v Cooper* 2012), with more than 90 per cent of convictions in both the United Kingdom and the United States obtained via guilty pleas (see Nobles and Schiff 2019: 102). Some have suggested that were all those defendants to carry their cases to trial rather than pleading guilty, the criminal justice system would be in danger of collapse (see, eg, *R v Caley and others* 2012: 6, per Lord Justice Hughes).

Despite its widespread use, the practice of reducing sentences for guilty pleas raises a number of concerns. Some of these concerns are epistemic: the existence of these incentives to plead guilty may undermine our confidence that a guilty plea constitutes sufficiently reliable evidence of a defendant's guilt (especially evidence of guilt beyond a reasonable doubt). Other concerns relate to principles commonly endorsed as constraints on sentencing. For example, giving those who plead guilty a reduced sentence may seem to violate cardinal proportionality insofar as their sentence is less severe than they deserve given the seriousness of their crime and their culpability for it. Also, these sentence reductions may seem to violate ordinal

*I presented a previous draft of this paper at a workshop at Worcester College, Oxford, in March 2022. I am grateful for feedback from the participants at that event, and in particular to Antony Duff and Jesper Ryberg for helpful discussion at that event and afterward.

proportionality insofar as equally culpable offenders guilty of similarly serious crimes may receive sentences of different severity when one but not the other pleads guilty (but see Chapter 3).

This chapter focuses on another worry about sentence reductions for guilty pleas, namely that these reductions violate the commonly accepted prohibition on punishment of the innocent. It is well documented that innocent defendants sometimes plead guilty. Rebecca Helm reports, for example, that in England and Wales from 2012 to 2019, 'of the 128 cases referred to the Court of Appeal by the Criminal Cases Review Commission as potential miscarriages of justice (for review of conviction or conviction and sentence) ... approximately 50 cases involved defendants who initially pleaded guilty' (Helm 2019: 425). In the United States, of 129 exonerations in 2020, 29 involved convictions resulting from guilty pleas (National Registry of Exonerations 2020). Innocence Canada has secured 24 exonerations since 1993, and at least five of these cases involved guilty pleas.¹

Innocent defendants may be motivated to plead guilty for a variety of reasons. They may not be able to afford the costs of going to trial. They may want to avoid staying in prison on remand until the trial begins. They may mistakenly believe they committed the act(s) with which they are charged. They may be seeking to protect the actual culprit from prosecution and punishment (see Sherrin 2011: 7–13). Arguably the most significant incentive for innocent defendants to plead guilty, however, is the desire to minimise their sentences. As noted before, the explicit aim of sentence reductions is to induce *guilty* defendants to plead guilty. But insofar as there is a prudential incentive not to risk the significantly more severe sentence that might follow a trial, this is an incentive to innocent defendants as well as guilty ones. Whether innocent or guilty, if a defendant believes it likely that he would be convicted if he took his case to trial anyway, then it will be tempting to secure a less severe sentence by pleading guilty. Even if the defendant believes it fairly likely that he can win if he takes his case to trial, the range of sentences he would face if convicted at trial might be so unacceptable to him that he might choose to plead guilty and secure the lesser sentence rather than even to risk the more severe sentence. Empirical research generally supports the claim that the prospect of a reduced sentence sometimes motivates innocent people to plead guilty (see, eg, Redlich et al 2017: esp. 349).²

This chapter examines whether the practice of offering sentence reductions for guilty pleas violates the prohibition on punishment of the innocent. First, in

¹ See the Innocence Canada website at <http://innocencecanada.com/exonerations>. For two of the 24 cases, the website did not include details of whether the original conviction resulted from a guilty plea or a trial.

² Some studies have found that innocent defendants are less likely to plead guilty than guilty defendants. (See, eg, Gazal-Ayal and Tor 2012). We can acknowledge this discrepancy in the prevalence of guilty pleas while still recognising that innocent defendants are often motivated to plead guilty by the prospect of sentence reductions.

Section II, I consider various interpretations of what this prohibition requires of a system of punishment. Next, in Section III, I contend that insofar as sentence reductions often provide significant prudential incentives to innocent people to plead guilty, these reductions run afoul of the most plausible construal of the prohibition on punishment of the innocent. In Section IV, I consider and respond to some potential objections.

II. The Prohibition on Punishment of the Innocent

The prohibition on punishment of the innocent, as standardly articulated, is essentially a retributivist constraint: it holds that it is inherently wrong to punish the innocent – even if such punishment would contribute to some valuable social end(s) – because this treatment would be *undeserved* by virtue of their prior conduct. To punish an innocent person is thus to wrong that person, to violate her rights. Although the principle is itself retributivist insofar as it is grounded in considerations of desert, one could endorse the constraint without endorsing a uniformly retributivist theory of punishment. Indeed, hybrid accounts of punishment typically adopt it as a constraining principle, along with the prohibition of disproportionate punishment of the guilty. In fact, it is often thought to be a decisive strike against pure consequentialist accounts that they cannot provide principled grounding for these constraints; at best, such theories can endorse the constraints contingently, insofar as the constraints promote the best consequences overall (see, eg, McCloskey 1957; Boonin 2008: 41–46).

Although the prohibition on punishment of the innocent is widely endorsed by punishment theorists, it is often insufficiently clear what precisely the constraint requires. We might construe it as a necessary condition of the justification of particular impositions of punishment: thus, a specific instance of punishment is not justified if the person punished is innocent. This seems plausible enough. (We might, of course, think the State could be excused, albeit not justified, in punishing an innocent person, if it did so mistakenly and having followed the appropriate procedures and observing the relevant safeguards.) But what if we focus instead on the institution of punishment rather than individual instances? What does the prohibition on punishment of the innocent tell us about the institution? Here I consider several possible interpretations of the prohibition.

First, we might interpret it as holding that for the institution of punishment to be justified, innocent people must never be punished. Richard Lempert endorses this sort of construal as the basis for his objection to capital punishment. He writes, ‘We know as a statistical matter that if a state executes often enough some innocent lives will be lost. ... Nothing about retributivism allows us to sacrifice the lives of unknown innocents in the interest of just vengeance’ (Lempert 1981: 1183). Although Lempert focuses on capital punishment, his argument has implications for punishment more generally. He contends essentially that any criminal

justice system will inevitably make mistakes, and thus it will subject at least some innocent people to burdensome treatment that they do not deserve.³ The retributivist constraint, as he interprets it, does not permit the imposition of undeserved burdens on the innocent even in rare cases, and so actual systems of punishment cannot be justified on retributivist terms.

But this seems too strict a construal. Even with robust safeguards in place to avoid conviction and punishment of innocent people, it is foreseeable that at least some innocent people will be mistakenly convicted and punished. To claim that the institution of punishment can only be justified if no innocent person is ever punished is, essentially, to claim that the institution of punishment cannot be justified. But those who endorse the prohibition on punishment of the innocent as a constraint are not typically abolitionists about punishment. This is not because they are unaware that even carefully designed systems for determining criminal guilt will inevitably result in the punishment of at least some innocent people. More plausibly, it is because they believe there is a compelling rationale for the institution of punishment (be it crime reduction, the communication of deserved censure, or something else), and they regard this rationale as sufficiently weighty that an institution that serves this function can be justified even if it foreseeably results in the inadvertent punishment of some innocent people. My own view is that the compelling rationale for punishment is to help protect public safety and well-being, and that this function is sufficiently important that the institution is not made illegitimate by the mistaken punishment of some innocent person(s). Thus, I suggest we should reject this strict interpretation of the prohibition on punishment of the innocent.

A second possible interpretation of this constraint is that it prohibits *deliberately* punishing the innocent (see, eg, Moore 1997: 158). Setting aside definitional questions of whether deliberately inflicting such treatment on those we know to be innocent constitutes punishment at all, it is clear at least that this construal of the prohibition does not imply that a system of punishment is unjustified if innocent people are ever mistakenly punished. If the first interpretation of the constraint was too strong, however, this second interpretation is too weak. It would fail to rule out a system in which, although innocent people were never deliberately punished, the risks of mistaken punishment of the innocent were widely recognised and entirely disregarded. That this weaker interpretation of the constraint cannot rule out a system of punishment that makes no effort to avoid inadvertent punishment of the innocent is, I suggest, good reason to reject it.

A third possible interpretation of the constraint is that it prohibits punishing the innocent except insofar as doing so facilitates punishment of the guilty to an

³ Lempert suggests that his conclusion about capital punishment does not generalise to other forms of punishment, specifically to life imprisonment, because many who are sentenced to life in prison 'will prove their blamelessness after conviction, be released early from prison, and in all probability be compensated to some extent' (Lempert 1981: 1183 n 16). For a persuasive rebuttal of Lempert on this point, see Alexander (1983: 235).

extent that the justice of the latter outweighs the injustice of the former. Larry Alexander describes this view as 'strong retributivism', and he contends that it offers a plausible way of resisting objections to retributivism grounded in concerns about punishment of the innocent. Alexander writes:

[Strong retributivism] deems as equally unjust failing to punish the guilty and punishing the innocent. Justice requires that we give everyone what he deserves, positively or negatively. A system of punishment will result in some innocent persons not getting what they deserve (freedom from punishment). Lack of a system of punishment will result in some guilty persons not getting what they deserve (punishment). Both results are equally unjust. But if unjust results are inevitable whether we punish or not, our preference is for fewer unjust results. A system of punishment, if well-designed, will produce fewer instances of injustice (1983: 238)

On this interpretation, a legal system that resulted in the punishment of some innocent people would be justifiable as long as such impositions of punishment made possible (or were an unavoidable side-effect of) the punishment of a greater number of guilty people who would not otherwise have been punished. A key problem with this interpretation, however, is that it equates the wrongness of failing to punish one offender with the wrongness of punishing one non-offender. Thus it leads to counterintuitive results. For example, if we have four suspects in custody for some crime, and we know that two of them are guilty and two are innocent, but we have no way of knowing which two are guilty, then according to 'strong retributivism' as Alexander characterises it, we would be as justified in punishing all four people as in letting all four go free. After all, letting all four go free would result in two cases of guilty people not getting what they deserve (punishment) whereas punishing all four would result in two cases of innocent people not getting what they deserve (freedom from punishment). Essentially, we could justifiably decide whether to punish none or all four with a coin flip. This will strike many of us as deeply unjust, though, because we think it is morally worse to subject an innocent person to intentionally burdensome, intentionally condemnatory treatment than to fail to subject a guilty person to such treatment.

This asymmetry in our intuitions about punishing the innocent and failing to punish the guilty is evident in William Blackstone's commonly endorsed principle that 'it is better that ten guilty persons escape than that one innocent suffer' (Blackstone 1893: ch 27). It is reflected in the common endorsement, by theorists who defend hybrid theories of punishment, of the constraint on punishment of the innocent but not a constraint on failure to punish the guilty (see, eg, Hart 1960; Corlett 2001, 78; Flanders 2017). It is also reflected in commonly held views about proportionality in sentencing: many argue that punishing offenders less severely than they deserve could in some cases be justified (for example, when mercy is warranted), but that punishing more severely than is warranted is never justified (see, eg, Duff 2001: 140–41; Corlett 2001: 78). The underlying intuition reflected in this asymmetry regarding deviations from proportionality in sentence

severity is essentially the same as the intuition underlying the asymmetrical views about punishment of the innocent and non-punishment of the guilty: it is worse to impose onerous, condemnatory treatment when it is not deserved than it is to fail to impose such treatment when it is deserved. Thus we should reject this interpretation, grounded in Alexander's 'strong retributivism', of the prohibition on punishment of the innocent.

It seems, then, that a requirement that a system of punishment, to be legitimate, must never mistakenly punish anyone is too strict. But also, to require only that the system not intentionally punish the innocent could permit a system in which large numbers of innocent people are inadvertently punished. Similarly, a prohibition on mistaken punishment of the innocent except insofar as this contributes to more instances of punishment of the guilty could be consistent with a substantial number of mistaken punishments of the innocent.

A fourth option, then, would be to interpret the prohibition on punishment of the innocent as holding that a system of punishment is unjustified if it results in significant numbers of innocent people being punished. One challenge presented by this interpretation of the constraint would be in determining how many instances of punishment of the innocent would rise to the level of constituting *significant numbers* and thus would render the system illegitimate. There seems no obvious, non-arbitrary basis for settling this issue. This is complicated by the fact that we do not know how many innocent people have been mistakenly punished. Estimates of rates of mistaken conviction vary significantly: in the US context, for example, estimates range from as low as 0.5 per cent to as high as 15 per cent (see, eg, Zalman 2012; Gross et al 2014; Walsh et al 2017; McCloskey 1989; Poveda 2001). So as a practical matter, interpreting the prohibition on punishment of the innocent as a requirement that mistaken punishment of the innocent remains below some given level will create challenges. More fundamentally, this construal of the prohibition would permit a penal system in which the State made no real effort to avoid punishment of the innocent but, by sheer luck, did not in fact punish many innocent people. Many of us would object to such a system, and rightly so, as it fails to take the risk of mistaken punishment of the innocent sufficiently seriously.

A more plausible account, I suggest, will accept that punishment of the innocent, even if inadvertent, represents a serious challenge to the legitimacy of a penal system, and that it is not something that can be legitimated merely by corresponding gains in punishment of the guilty. But it also will acknowledge that the State has a compelling interest in protecting public safety and well-being, and that insofar as the institution of punishment plays a key role in this regard, the institution will not be rendered unjustified by the mistaken conviction and punishment of a comparatively small number of innocent people.

My suggestion for such a principle draws from another context, namely philosophical debates about the ethics of war. A central question in these debates is when, if ever, causing inadvertent civilian casualties can be morally permissible in the pursuit of some valuable goal in times of war. One prominent answer to

this question is to invoke the Doctrine of Double Effect, which holds, basically, that if the casualties are an unintended side-effect of an action, rather than the means by which the action achieves its end, and if the good outcomes promoted by the action sufficiently outweigh the costs, then the action can be justified (see, eg, McIntyre 2018). The Doctrine of Double Effect has proved controversial (see Bennett 1981; see also Fitzpatrick 2012) and my aim here is not to defend it. Rather, I want to focus on one prominent objection to this doctrine, namely that it is too permissive. It is not enough that combatants do not intend to harm civilians; rather, what is required is that combatants *intend not to harm them*. Michael Walzer thus proposed that military actors must aim to avoid the foreseeable harm by making a positive commitment to save innocent lives, and indeed that they must be willing to accept additional risks or sacrifice some benefits in order to avoid harm to innocents (Walzer 1977: 151–59; see also Lee 2004).

In my view, these considerations are relevant to State punishment as well. For State impositions of punishment to be legitimate, it is not enough that the State not intend to punish innocent people; it should make a serious commitment not to punish the innocent. Institutions can reflect this intention by taking significant steps – even accepting some additional risks or sacrificing some benefits – to avoid punishment of the innocent. This is what the prohibition on punishment of the innocent requires of legitimate systems of punishment. It does not require that no innocent person ever be punished, but it requires more than merely the avoidance of deliberate punishment of the innocent. And it does not treat punishment of the innocent as a wrong that can easily be traded off against the supposedly equal wrong of non-punishment of the guilty. Rather, punishment of the innocent is a significantly more serious wrong than non-punishment of the guilty, and although the compelling interest in maintaining an institution of State punishment to help protect public safety and well-being is sufficiently significant that the institution's legitimacy is not cancelled by some instances of inadvertent punishment of the innocent, the institution's legitimacy does require that the State demonstrate its commitment not to punish the innocent by taking substantial steps to avoid this. In the next section, I consider whether this requirement is met in current legal practice.

III. Criminal Trials, Guilty Pleas and Not Punishing the Innocent

Do our criminal justice practices reflect a serious commitment not to punish the innocent? In this section, I contrast criminal trials with the plea process. In criminal trials, the State calls on defendants to answer to charges of wrongdoing, and it seeks to determine whether they are guilty because there is a compelling public interest in holding accountable those who have committed crimes. But also, a

host of protections are in place to help guard against the conviction of innocent defendants. The presumption is that defendants are innocent and the burden falls on the prosecution to prove their guilt, rather than requiring defendants to prove their innocence. The prosecution must prove guilt to a very high standard, beyond reasonable doubt (rather than, say, the lower standards of ‘balance of probabilities’ or ‘preponderance of evidence’ used in civil cases). Defendants typically have a right to face, and have their defence counsel cross-examine, their accusers. They also have a right to access any exculpatory evidence possessed by the prosecution. Double jeopardy protections prevent those who are found not guilty from being prosecuted repeatedly for the same crime. What is more, defendants who are found guilty at trial have a right to appeal their convictions. Thus, while criminal trials reflect the State’s interest in determining whether individuals are guilty of the crimes of which they are accused, so that the guilty can be held accountable, the various protections afforded to defendants also reflect a serious commitment to protect against conviction and punishment of innocent people.

As discussed earlier, however, the lion’s share of convictions in the United Kingdom, the United States and many other countries come as a result of guilty pleas, typically incentivised by sentence reductions (see Fair Trials 2022). Criminal justice today is, as Justice Kennedy pointed out (*Lafler v Cooper* 2012), more a system of pleas than of trials. We should focus, then, on the plea system and whether it reflects a serious commitment to avoid punishment of innocent people.

Most importantly, as I have discussed, a system in which sentence reductions attach to guilty pleas creates a prudential incentive for defendants, innocent as well as guilty, to plead guilty. Arguably, a system that creates such incentives cannot be said to reflect a serious commitment to avoid punishment of the innocent. But several further features of existing plea systems exacerbate the issue. For example, the prospect of innocent defendants being induced to plead guilty is that much greater in jurisdictions where prosecutors intentionally offer the largest sentence reductions when their cases are the weakest. A Chicago defence attorney describes it this way:

When a prosecutor has a dead-bang case, he is likely to come up with an impossible offer like thirty to fifty years. When the case has a hole in it, however, the prosecutor may scale the offer all the way down to probation. The prosecutors’ goal is to get something from every defendant, and the correctional treatment the defendant may require is the last thing on their minds. (Quoted in Alschuler 1968: 60; see also, eg, Scott and Stuntz 1992: 1946)

Weakness of a case is not always a symptom of a defendant’s innocence, but often it is. Thus, as Albert Alschuler writes, ‘the greatest pressures to plead guilty are brought to bear on defendants who may be innocent’ (1968).

Another feature of many plea systems is that the right of access to exculpatory evidence does not apply at the pre-plea stage. The prosecution is only required to disclose evidence that might indicate a defendant’s innocence once the defendant

has pleaded not guilty.⁴ Thus whereas a defendant who takes his case to trial has a legal right to see exculpatory evidence police investigators may have found, it is often the case that defendants considering whether to plead guilty and receive a sentence reduction or go to trial and lose the reduction are not afforded access to this information.

The plea process also differs from the trial process with respect to the likelihood of appeals. In England and Wales, the Court of Appeal treats convictions based on guilty pleas as presumptively safe. As the court explained in *Asiedu*,

ordinarily, once he has admitted such facts by an unambiguous and deliberately intended plea of guilty, there cannot then be an appeal against his conviction, for the simple reason that there is nothing unsafe about a conviction based on the defendant's own voluntary confession in open court. (*R v Asiedu* 2015: 19)

The right to appeal a conviction is therefore much more limited for those who plead guilty than for those found guilty at trial, and successful guilty plea appeals are rare. As Richard Nobles and David Schiff write:

A plea will be treated as a nullity if it can be shown to be either equivocal or involuntary. Where it is not, it will be treated by the Court of Appeal as an acknowledgement of guilt by the defendant, giving rise to a definite presumption that the conviction is safe. That presumption operates in a manner close to a bar on appeals, given the difficulties of rebutting it. (2019: 105; see also Campbell et al 2019: 315–16)

So whereas an innocent defendant who loses his case at trial still has a reasonable chance to appeal his conviction – grounds for appeal might include fresh evidence, failure to disclose exculpatory evidence, and jury irregularities – an innocent defendant who pleads guilty has comparatively little prospect of a successful appeal.

Thus with the plea process we have a system in which defendants are offered significant incentives in the form of sentence reductions to plead guilty (in some cases, larger incentives the weaker the case against them), in which they are not afforded access to exculpatory information, and in which they have relatively little chance to appeal their conviction later. In my view, such a process does not reflect a serious commitment to avoid conviction and punishment of the innocent. I contend, then, that the plea process as it currently operates is inconsistent with the most plausible construal of the prohibition on punishment of the innocent.

Could current plea practices be reformed so as to be consistent with the prohibition on punishment of the innocent? Defendants could be given access to exculpatory evidence prior to making their plea decisions. Also, those who plead

⁴In the United Kingdom, the duty to disclose exculpatory evidence commences once the defendant (in the Magistrates' Court) has pleaded not guilty and the case goes to trial, or (for cases going to the Crown Court) once the case has been sent to trial (see Ministry of Justice 2020: 16). In the United States, federal courts have split about whether prosecutors have a duty to disclose prior to a guilty plea or not until the case goes to trial (see Petegorsky 2013; Clarke 2019).

guilty could be afforded more opportunity to appeal their convictions, in line with those who are found guilty at trial. And in those legal systems where prosecutors have discretion to offer the largest sentence reductions when their cases are the weakest, this discretion could be eliminated. Still, as long as a system offers innocent defendants a prudential incentive, in the form of a significant sentence reduction, to plead guilty, it is difficult to see the system as reflecting a serious commitment to avoid punishment of the innocent.

Rather than eliminating sentence reductions, of course, we might endorse reducing them to a level at which they would not provide innocent defendants with a significant incentive to plead guilty. Liz Campbell, Andrew Ashworth and Mike Redmayne, despite scepticism about the legitimacy of offering sentence reductions at all, suggest that sentence reductions of no more than 10 per cent would reduce the incentive to plead guilty to a tolerable level and ensure that the innocent defendant's freedom of choice is maintained (Campbell et al 2019: 341).⁵ The worry with such a proposal, however, is that it either will not address the injustice it aims to address, or it will undermine the central rationale for offering sentence reductions for guilty pleas. That is, it seems that either the prospect of a sentence reduction will provide a compelling prudential reason to plead guilty for a defendant worried about the prospect of losing at trial and facing a stiffer sentence, or it will not. If this prospect does provide such a prudential reason for guilty defendants, then it will, by the same token, be a reason for many innocent defendants too. If it does not provide such a reason for innocent defendants (because we have reduced the maximum discount to a level that would not be a significant incentive for them), then it will not provide such a reason for guilty defendants either, and thus we would lose the central rationale for the discounts: to persuade guilty defendants to plead guilty and thus forego a trial.

It is possible, however, despite the worry just discussed, that some modest level of sentence discount exists that would entice a significant number of guilty defendants to plead guilty without similarly persuading a significant number of innocent defendants to plead guilty. Perhaps innocent defendants are, on balance, somewhat less receptive to the prudential incentives of a sentence discount, as they are also motivated (whereas guilty defendants are not) by the knowledge that they are innocent and a desire not to confess to a crime they did not commit (see, eg, Gazal-Ayal et al 2012). It is an empirical question, beyond the scope of this chapter, whether some modest level of sentence discount could serve to motivate guilty pleas from a significant number of guilty defendants while not also motivating guilty pleas from a significant number of innocent defendants. For present purposes, however, two things are worth noting. First, if such a modest level of sentence reduction could be found, it would be significantly lower than the sentence discounts currently on offer in many legal systems.

⁵ See also Gormley et al (2020: 14) for the suggestion that current levels of discounts, at least in England and Wales, may often be sufficiently modest in this respect.

Second, even if some modest level of sentence discount could be said not to incentivise guilty pleas from a significant number of innocent defendants, reducing existing guilty plea sentence reductions to such a level would not in itself be sufficient to render the plea process consistent with the prohibition on punishment of the innocent. I have contended that the most plausible interpretation of this constraint is that the State should demonstrate a serious commitment to avoid punishment of the innocent. I have further argued that a system that creates incentives for innocent people to plead guilty violates this constraint. But even if the State does not actively create such incentives, this in itself is not enough to demonstrate its commitment to avoid punishment of the innocent. Analogously, it is not sufficient to demonstrate my commitment to minimising vandalism of university property that I stop offering cash prizes to my students to see who can hit the vice chancellor's window with the most tomatoes in under a minute. More generally, not incentivising X is not sufficient to demonstrate a commitment to minimising X. Thus, in addition to eliminating sentence reductions, or at least reducing them to a level that does not provide a significant incentive for innocent defendants to plead guilty, the State would also need to incorporate measures into the plea process to help ensure that innocent defendants do not plead guilty. As noted earlier, such measures should include, at least, a requirement to disclose to defendants exculpatory evidence prior to their plea decision, and more robust opportunities to appeal convictions resulting from guilty pleas. In the next section, I consider potential objections to my account.

IV. Objections and Responses

I have argued that a commitment to avoid punishment of the innocent prohibits creating incentives to plead guilty, in the form of sentence discounts, because, even if these discounts are expressly intended to motivate guilty defendants to plead guilty, we know they also serve to motivate innocent defendants to plead guilty. One might object, however, that we create analogous incentive structures in other contexts without apparently doing anything impermissible. Consider, for example, the sale of alcohol and assume for argument's sake that whereas it is permissible to sell alcohol to adults, we should not sell to minors. Suppose, too, that sellers not only should not deliberately sell to minors but, further, should demonstrate a commitment not to sell to them. Now suppose that a store manager launches a promotional campaign to boost sales, including the offer of some incentive (eg, entry into a prize draw with each purchase). The aim is to motivate adults to buy, but the store manager knows the new incentive will also motivate at least some minors to attempt to purchase alcohol. Despite introducing some additional safeguards (more stringent ID checks, etc), the manager is reasonably sure that at least some clever minors, motivated by the new incentive, will manage to buy alcohol. The objection, then, is that in such a case it is nevertheless permissible to offer the

incentive, and as this case is analogous in relevant respects to the offering of guilty plea sentence discounts, it is similarly permissible to offer such discounts.⁶

There is an important respect, however, in which we can distinguish the offer of guilty plea sentence discounts from cases such as the example of incentivising the purchase of alcohol. Although we think it is undesirable for various reasons for minors to purchase and drink alcohol, the minor who buys and drinks alcohol is not thereby wronged. Thus, when the store manager mistakenly sells alcohol to the minor, he does not wrong the minor (or contribute to the minor's being wronged).⁷ By contrast, the State does wrong the innocent person when it punishes her. I mentioned earlier that the State may be excused for punishing the innocent person if, say, it took appropriate steps to determine her guilt and afforded her proper safeguards, etc. But the innocent person subjected to intentionally burdensome, condemnatory treatment is wronged nonetheless. I have argued above that the mere fact that the State inadvertently wrongs some individuals does not in itself demonstrate that the institution of punishment is unjustified. But for the institution to be justified, it must be structured in a way that demonstrates a serious commitment to avoid punishing the innocent. This means, among other things, that the State should not create incentives for innocent people to plead guilty.

Does a commitment to avoid punishment of the innocent also prohibit the State from making defendants aware of existing reasons to plead guilty? A prosecutor, acting as the agent of the State, might explain to a defendant the strength of the case against her, and on the basis of this information the defendant might decide to plead guilty (even if there were no guilty plea sentence reductions, the defendant might simply not want to go to trial if an acquittal seemed highly unlikely). Such information might serve to motivate innocent defendants as well as guilty ones. In this sort of case, however, the prosecutor does not create incentives to plead guilty; he merely makes defendants aware of existing considerations that may be relevant to their choice of plea. The likelihood of losing at trial may be an incentive to some defendants to plead guilty and avoid the hassle, but the prosecutor does not create this incentive. In my view, making defendants aware of the strength of the case against them is consistent with demonstrating a commitment to avoid punishment of the innocent.

A different sort of objection to my critique of guilty plea sentence reductions might focus on the benefits of the practice. As mentioned before, guilty pleas may relieve victims and witnesses of the stress of having to testify, and they may help reduce the impact of the crime on victims. If sentence reductions facilitate these benefits, then this seems to count in their favour. What is more, insofar as sentence reductions result in a higher proportion of guilty defendants being convicted and

⁶Thanks to Antony Duff for suggesting this example.

⁷One might argue, I suppose, that the State does wrong minors when it (even inadvertently) sells them alcohol. I am sceptical that such an argument would succeed. But if it were successful, then it would follow on my account that creating incentives to purchase alcohol, which will inevitably motivate minors as well as adults, would be impermissible.

punished than would otherwise be the case, they would seem, at least in this regard, to further the ends of maintaining a system of punishment. This is true whether the appropriate rationale for punishment is to help reduce crime (through deterrence, incapacitation or reform) or to convey deserved censure or mete out deserved suffering. One might worry, then, that my critique of sentence reductions does not take seriously enough the benefits that guilty pleas promote, and thus the importance of sentence reductions in achieving these valuable ends.

I accept that there are benefits to allowing defendants the option of pleading guilty rather than taking their case to trial. It is inappropriate, however, for the State to pursue these benefits at the expense of innocent defendants. This does not mean that offering the option of guilty pleas is unjustified if it ever results in innocent defendants pleading guilty and facing punishment. But I have argued that the State must demonstrate a serious concern to avoid punishment of the innocent. Demonstrating such a concern need not commit the State to removing the option of guilty pleas, and thus requiring that every case be tried in court. But it does require, among other things, that the State not create significant prudential incentives to innocent defendants to plead guilty. If not creating such incentives means also that more guilty defendants take their cases to trial, and this leads to additional stress for some victims or witnesses, or if it means that guilty defendants are acquitted who would otherwise have pleaded guilty and faced punishment, then I acknowledge that these are genuine costs. But in my view, they are outweighed by the wrong done when innocent defendants are incentivised to plead guilty.

A different possible objection focuses on the value of guilty plea sentence reductions to innocent defendants themselves. According to this objection, if our concern is to protect against punishment of the innocent, then this could count in favour of guilty plea sentence discounts as well as against them. First, we know that some number of innocent defendants will plead guilty even without the incentive of sentence reductions. If sentence reductions are eliminated, then these innocent people will be punished more severely than they will be if the reductions are offered (see Chapter 3). Similarly, some number of innocent defendants may, without the incentive of a guilty plea sentence reduction, choose to take their case to court, and some of these defendants will be convicted. These defendants might have chosen to plead guilty to get the sentence reduction had it been offered; thus here again, the availability of the sentence reduction might contribute to less punishment of the innocent.

I have contended that the best construal of the prohibition on punishment of the innocent is that the State is required to demonstrate a serious commitment to avoid such punishment. Suppose, then, a system in which guilty plea sentence reductions are not offered; suppose, too, that the State knows that in this system some number of innocent defendants plead guilty and are subject to punishment, and also that some number of innocent people take their cases to court and are convicted. The question is what the State might do to demonstrate a serious commitment to avoid punishing the innocent. Specifically, could the State demonstrate such a commitment by introducing a practice of offering sentence

discounts as prudential incentives to plead guilty, knowing that this will persuade some innocent defendants to plead guilty who would not otherwise have done so but also will lessen the punishment of those people who were going to plead guilty regardless or who would otherwise take their cases to court and lose?

It strikes me as counterintuitive that introducing such a practice would demonstrate a commitment to avoid punishing the innocent. To see why, consider again the prohibition on punishment of the innocent. Earlier, I wrote that this constraint is essentially a retributivist principle (though it may be an element of hybrid accounts as well as pure retributivist accounts); it holds that punishment of the innocent is inherently wrong because it is undeserved. Although retributivism is typically contrasted with consequentialism, some theorists have pointed out that one might endorse a consequentialist version of retributivism, according to which just desert, which has inherent value, is to be promoted or maximised (see, eg, Moore 1997: 156–58; but see Dolinko 1997). On a consequentialist view, whatever it is that we take to have value, the appropriate response is to promote it (see Pettit 1991: 19). Importantly, if violating or undermining the value in some instances leads to more of it overall, then this is what consequentialism tells us to do.⁸ If we interpret the retributivist prohibition on punishment of the innocent in this consequentialist sense, then it might make sense to institute a practice that incentivises some innocent defendants to plead guilty and face undeserved punishment if doing so contributed to less undeserved punishment more generally.

The problem with interpreting the constraint as a consequentialist-retributivist principle is that it opens the constraint up to just the sort of objection that retributivists have often levelled against consequentialist theories of punishment more generally: this is the Kantian objection that consequentialist punishment uses the offender as a mere means to promote some ostensibly valuable social end (see, eg, Murphy 1973: 219). To attempt to justify a practice that incentivises some innocent defendants to plead guilty and face punishment (which thereby wrongs them) on grounds that this practice promotes less punishment of the innocent overall appears similarly to use these innocent defendants as mere means, or tools, to the end of minimising punishment of the innocent. It is true that in some cases innocent defendants who are persuaded to plead guilty to secure the reduced sentence would otherwise have been convicted at trial; it might seem that these defendants are not used as *mere* means in the service of reducing punishment of the innocent, as they themselves benefit from such a reduction. But there will be some number of innocent defendants who will be persuaded to plead guilty by the prudential

⁸This is the case not only for direct forms of consequentialism such as act consequentialism, but also for indirect forms of consequentialism such as rule consequentialism or motive consequentialism. On any of these accounts, the rightness or wrongness of our action, rules for action, motives for action, etc, will be a matter of their consequences. And if occasionally acting in violation of what we value (or acting on a rule that tells us sometimes to violate what we value, or acting from a motive to violate what we value) thereby promotes more of the value overall, then consequentialism tells us to do that.

incentive of a sentence discount who would have been acquitted had they taken their case to trial.⁹ These people, it seems, are treated as mere means in the service of reducing punishment of the guilty overall. But again, retributivists often critique consequentialist accounts specifically for implying that it can be permissible to use some people as mere means to securing some valuable social end.

I contend that the prohibition on punishment of the innocent is best interpreted instead as a deontological-retributivist principle rather than a consequentialist-retributivist principle. A deontological account holds that the right response to whatever we value may sometimes be to respect it, or honour it, rather than to promote it (see Pettit 1991: 19). For a deontologist, respecting some value will sometimes mean not violating it even if doing so would result in an overall net gain of that value. (A deontologist who values honesty, for example, may hold that lying is wrong even if by telling a few lies a person could promote more honesty overall.) A deontological interpretation of the prohibition on punishment of the innocent holds that such punishment is unjustified insofar as it is undeserved, and the State should demonstrate a serious commitment to avoid such punishment; this is so even if a practice that incentivises some innocent people to plead guilty contributes to less punishment of the innocent overall.

This same sort of reasoning can also help to make clear why it would not be appropriate to impose a sentence *premium* for guilty pleas. One might think, given the account I have defended here, with its emphasis on demonstrating a commitment not to punish the innocent, that the State should therefore impose a premium for guilty pleas to disincentivise them. Such a suggestion will likely strike many as implausible, but one might wonder whether an endorsement of not offering guilty plea sentence discounts must also commit us to this further measure.¹⁰

Demonstrating a commitment to avoid punishment of the innocent need not commit us to take every step that might serve this end, however; we can acknowledge that there are appropriate and inappropriate ways to demonstrate this commitment. One key problem with imposing guilty plea sentence premiums is that guilty defendants who want to accept responsibility for what they have done by pleading guilty will be punished more severely for doing so. In effect, this would be to subject such defendants to a greater degree of intentionally burdensome, censoring treatment simply because they chose to be honest, to take responsibility for what they did. The additional punishment represented by the sentence premium would thus be undeserved, and just as punishing innocent people wrongs them insofar as the punishment is undeserved, punishing guilty people more severely than they deserve wrongs them, too. If, as I have argued, the concern motivating the prohibition on punishment of the innocent is the deontological-retributivist concern not to subject people to harsh treatment they do not deserve, then we cannot justify a policy that leads to the imposition of undeserved degrees

⁹We should expect such cases to be especially common in jurisdictions where, as discussed earlier, prosecutors offer the largest discounts when their cases are the weakest.

¹⁰Thanks to Jesper Ryberg for raising this concern.

of punishment on guilty defendants simply for their choice to own up to what they did, even if such a policy also serves to reduce undeserved punishment of some innocent defendants by dissuading them from pleading guilty. Subjecting honest guilty defendants to this undeserved punishment (thus wronging them) merely in the service of reducing undeserved punishment of other, innocent defendants would be to use the guilty defendants as mere means to this further end.

There are other ways to demonstrate a commitment to avoid punishment of the innocent, however, that do not require using some defendants as mere means. For those innocent people who plead guilty regardless of the sentence reduction, we should ask why this happens and if there are any steps that we can take to minimise this. Some innocent defendants plead guilty to avoid long periods of pre-trial detention. We should thus work to shorten the time between arrest and the trial so that this becomes less of an incentive for innocent people to plead guilty. Other defendants plead guilty because they cannot afford the financial costs of a trial. A concern to avoid punishment of the innocent should spur us to look for ways to alleviate the financial burdens of taking one's case to trial. For those innocent defendants who do take their cases to trial, there are numerous measures in place (some of which I mentioned earlier) to protect against conviction and punishment of the innocent. Although, foreseeably, some innocent people are mistakenly convicted anyway, it is nonetheless the case that the State demonstrates a commitment to avoid such outcomes. If, however, we begin to discover that large numbers of people convicted at trial are in fact innocent, then a commitment to avoid punishment of the innocent should spur us to consider whether additional safeguards are warranted.

V. Conclusion

I have contended that the practice of offering significant sentence discounts as prudential incentives to defendants to plead guilty violates the prohibition on punishment of the innocent. In this concluding section, I want to return to the claim, mentioned at the start of this chapter, that were most defendants who now plead guilty to take their cases to trial instead, the criminal justice system would be in danger of collapse. Insofar as sentence discounts incentivise guilty pleas, one might worry that any proposal to eliminate or greatly restrict the use of such discounts would threaten to bring the system to a halt. To make matters worse, I have also suggested that a commitment to avoid punishing innocent people should lead us to require disclosure of exculpatory evidence to defendants at the pre-plea stage; to offer defendants convicted based on guilty pleas more opportunity to appeal their convictions; to reduce the time spent in pre-trial detention; and to reduce or eliminate the financial burdens of trials for defendants. These measures would all require additional resources. Thus it may seem that these proposals, albeit well intended, are ultimately unrealistic.

If the criminal justice system would creak under the weight of what I take to be reasonable proposals on their merits, then I suggest we would do better to ask why the system itself has got to this point. It has not always been the case that more than 90 per cent of convictions came from pleas rather than trials (see, eg, Langbein 2003, 18–20; Rakoff 2014; Chapter 12). Why is the current system such that without a practice that incentivises high levels of guilty pleas, it cannot function? One plausible contributing factor is that too much conduct is criminalised. In England and Wales, for example, the Law Commission reported that more than 3,000 crimes were added to the statute book between 1997 and 2010 (Criminal Liability in Regulatory Contexts 2010: 5; on over-criminalisation in the American context, see Husak 2008). Lower levels of criminalisation would lead to fewer people being brought into the criminal justice process, which would reduce the strain on the system's resources. Conversely, if the public or our lawmakers are unwilling to accept a smaller criminal justice system, then we should be prepared to pay the additional costs required to ensure that the system operates in a justifiable way. In particular, if demonstrating a serious commitment to avoid punishing the innocent means the criminal justice system becomes less efficient or more expensive, then that is the price we should pay. To borrow a quote from detective Kim Griggs in *The Wire*, 'sometimes things just got to play hard'.

References

- Alexander, L (1983) 'Retributivism and the Inadvertent Punishment of the Innocent' 2 *Law and Philosophy* 233.
- Alschuler, AW (1968) 'The Prosecutor's Role in Plea Bargaining' 36 *University of Chicago Law Review* 50.
- Bennett, J (1981) 'Morality and Consequences' in S McMurrin (ed), *The Tanner Lectures on Human Values* II (Salt Lake City, UT, University of Utah Press).
- Blackstone, W (1893) *Commentaries on the Laws of England in Four Books*, vol 2 (Philadelphia, PA: J.B. Lippincott Co) <https://oll.libertyfund.org/title/sharswood-commentaries-on-the-laws-of-england-in-four-books-vol-2>
- Boonin, D (2008) *The Problem of Punishment* (Cambridge, Cambridge University Press).
- Campbell, L, Ashworth, A and Redmayne, M (2019) *The Criminal Process*, 5th edn (Oxford, Oxford University Press).
- Clarke, E (2019) 'Entitling the Accused to Exculpatory Evidence: Why Prosecutors Should Have to Disclose During Plea Bargaining' 56 *American Criminal Law Review* 65.
- Corlett, JA (2001) 'Making Sense of Retributivism' 76 *Philosophy* 77.
- Criminal Liability in Regulatory Contexts (2010) *The Law Commission, Consultation Paper No. 195*, www.lawcom.gov.uk/app/uploads/2015/06/cp195_Criminal_Liability_consultation.pdf
- Dolinko, D (1997) 'Retributivism, Consequentialism, and the Intrinsic Goodness of Punishment' 16 *Law and Philosophy* 507.
- Duff, RA (2001) *Punishment, Communication, and Community* (Oxford, Oxford University Press).

- Fair Trials (2022) *The Disappearing Trial*, www.fairtrials.org/app/uploads/2022/01/The-Disappearing-Trial-report.pdf
- Fitzpatrick, WJ (2012) 'The Doctrine of Double Effect: Intention and Permissibility' 7 *Philosophy Compass* 183.
- Flanders, C (2017) 'Punishment, Liberalism, and Public Reason' 36 *Criminal Justice Ethics* 61.
- Gazal-Ayal, O and Tor, A (2012) 'The Innocence Effect' 62 *Duke Law Journal* 339.
- Gormley, J, Roberts, JV, Bild, J and Harris, L (2020) 'Sentence Reductions for Guilty Pleas: A Review of Policy, Practice, and Research', *Sentencing Academy*, <https://sentencingacademy.org.uk/wp-content/uploads/2020/12/Sentence-Reductions-for-Guilty-Pleas-3.pdf>
- Gross, SR, O'Brien, B, Hu, C and Kennedy, EH (2014) 'Rate of False Conviction of Criminal Defendants Who Are Sentenced to Death' 111 *Proceedings of the National Academy of Sciences* 7230.
- Hart, HLA (1960) 'Prolegomenon to the Principles of Punishment' 60 *Proceedings of the Aristotelian Society* 1.
- Helm, RK (2019) 'Constrained Waiver of Trial Rights? Incentives to Plead Guilty and the Right to a Fair Trial' 46 *Journal of Law and Society* 423.
- Husak, D (2008) *Overcriminalization: The Limits of the Criminal Law* (Oxford, Oxford University Press).
- Kelly, W, Hussemann, J, Flynn, A, Yahner, J and Golian, L (2017) 'Estimating the Prevalence of Wrongful Conviction', www.ncjrs.gov/pdffiles1/nij/grants/251115.pdf
- Langbein, JH (2003) *The Origins of Adversary Criminal Trial* (Oxford, Oxford University Press).
- Lee, S (2004) 'Double Effect, Double Intention, and Asymmetric Warfare' 3 *Journal of Military Ethics* 233.
- Lempert, RO (1981) 'Desert and Deterrence: An Assessment of the Moral Bases of the Case for Capital Punishment' 79 *Michigan Law Review* 1177.
- McCloskey, HJ (1957) 'An Examination of Restricted Utilitarianism' 66 *The Philosophical Review* 466.
- McCloskey, J (1989) 'Commentary: Convicting the Innocent' 8 *Criminal Justice Ethics* 2.
- McIntyre, A (2018) 'Doctrine of Double Effect' in EN Zalta (ed), *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/spr2019/entries/double-effect/>
- Ministry of Justice (2020) *Criminal Procedure and Investigations Act 1996 (sec 23(1)) Code of Practice* (September), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/931173/Criminal-procedure-and-investigations-act-1996.pdf
- Moore, M (1997) *Placing Blame* (Oxford, Oxford University Press).
- Murphy, JG (1973) 'Marxism and Retribution' 2 *Philosophy & Public Affairs* 217.
- National Registry of Exonerations (2020) 'Annual Report', www.law.umich.edu/special/exoneration/Documents/2021AnnualReport.pdf
- Nobles, R and Schiff, D (2019) 'Criminal Justice Unhinged: The Challenge of Guilty Pleas' 39 *Oxford Journal of Legal Studies* 100.
- Petegorsky, M (2013) 'Plea Bargaining in the Dark: The Duty to Disclose Exculpatory Brady Evidence During Plea Bargaining' 81 *Fordham Law Review* 3599.
- Pettit, P (1991) 'Consequentialism' in P Singer (ed), *A Companion to Ethics* (Oxford, Blackwell Publishers Ltd).
- Poveda, TG (2001) 'Estimating Wrongful Convictions' 18 *Justice Quarterly* 689.

- Rakoff, JS (2014) 'Why Innocent People Plead Guilty', *The New York Review of Books*, www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/
- Redlich, AD, Bibas, S, Edkins, VA and Madon, S (2017) 'The Psychology of Defendant Plea Decision Making' 72 *American Psychologist* 339.
- Scott, RE and Stuntz, WJ (1992) 'Plea Bargaining as a Contract' 101 *Yale Law Journal* 1909.
- Sentencing Council (2017) *Reduction in Sentence for Guilty Plea: Definitive Guideline*, www.sentencingcouncil.org.uk/wp-content/uploads/Reduction-in-Sentence-for-Guilty-Plea-definitive-guideline-SC-Web.pdf
- Sherrin, C (2011) 'Guilty Pleas from the Innocent' 30 *Windsor Review of Legal and Social Issues* 1.
- Walsh, K, Hussemann, J, Flynn, A, Yahner, J and Golian, L (2017) 'Estimating the Prevalence of Wrongful Conviction', US Department of Justice, www.ncjrs.gov/pdffiles1/nij/grants/251115.pdf.
- Walzer, M (1977) *Just and Unjust Wars* (New York, Basic Books).
- Zalman, M (2012) 'Qualitatively Estimating the Incidence of Wrongful Convictions' 48 *Criminal Law Bulletin* 221.

Cases

United Kingdom

R v Asiedu [2015] EWCA Crim 714

R v David Caley and others [2012] EWCA Crim 2821

United States

Lafler v Cooper 566 US 156 (2012)

