Public Reason and the Justification of Punishment

Zachary Hoskins

Submit your article to this journal ⧉

View related articles ⧉

View Crossmark data ⧉

Routledge
Taylor & Francis Group

ARTICLE

Check for updates
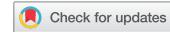
Public Reason and the Justification of Punishment

ZACHARY HOSKINS ⓘ

Chad Flanders has argued that retributivism is inconsistent with John Rawls's core notion of public reason, which sets out those considerations on which legitimate exercises of state power can be based. Flanders asserts that retributivism is grounded in claims about which people can reasonably disagree and are thus not suitable grounds for public policy. This essay contends that Rawls's notion of public reason does not provide a basis for rejecting retributivist justifications of punishment. I argue that Flanders's interpretation of public reason is too exclusionary: on it, public reason would rule out any prominent rationale for punishment. On what I contend is a better interpretation of public reason, whether retributivism would be ruled out as a rationale for punishment depends on whether a retributivist account can be constructed from shared political commitments in a liberal democracy. Some prominent versions of retributivism meet this requirement and so are consistent with public reason.

I. Introduction

The political philosopher John Rawls wrote comparatively little about the justification of state punishment.[1] In recent years, however, political and legal theorists have sought to draw on Rawlsian principles in various ways to glean insights about punishment.[2] These scholars are part of a broader trend in penal theory that regards questions related to punishment's justification as ultimately political questions. On this view, we should understand punishment as centrally an exercise of coercive power by the state, and we should then ask whether, when, for what reasons, of what severity, and in what mode such exercises of state power are legitimate. Not all attempts to grapple with punishment

Zachary Hoskins, Associate Professor, Department of Philosophy, University of Nottingham, UK. Email: Zachary. Hoskins@nottingham.ac.uk

as a problem for political theory have been of a Rawlsian flavor,[3] although it is perhaps understandable given the influence of Rawls's political theory generally that it has inspired so many accounts.

In recent work, Chad Flanders draws on Rawls's notion of public reason as the basis of an argument against retributivist justifications of punishment.[4] Flanders contends that retributivism as a rationale for punishment is inconsistent with Rawls's core notion of public reason, which sets out those considerations, drawn from the public political culture, on which legitimate exercises of state power can be based. Flanders argues that retributivism is grounded in claims about which people can reasonably disagree, and which are thus not suitable grounds for legitimate exercises of state power. He contends that a more suitable rationale for punishment, one consistent with the constraints of public reason, is to maintain a credible threat to deter offenders from reoffending (i.e. specific deterrence).

In this essay, I argue that Rawls's notion of public reason does not provide a clear basis for rejecting retributivist justifications of punishment.

In what follows, I first briefly explain Rawls's notion of public reason and then reconstruct how Flanders draws on this notion to ground his objection to retributivism. Then, I look more closely at what public reason requires, and I argue that Flanders's interpretation is too exclusionary: on it, public reason would rule out retributivism, but also any other rationale for punishment (as well as rationales for the abolition of punishment). On what I contend is a better interpretation of public reason, whether retributivism would be ruled out as a candidate rationale for punishment depends on whether a retributivist account can be constructed from the shared political commitments of a liberal democratic society. I argue that at least some prominent versions of retributivism meet this requirement and so are consistent with public reason. After addressing potential lines of objection, I conclude by briefly considering some broader implications of my argument for the ambition of appealing to Rawlsian public reason to adjudicate among competing theories of punishment, or competing accounts in political philosophy more generally.

II. Public Reason and Retributivism

A central question Rawls takes up in his later work is how exercises of political power can be legitimate in liberal democratic societies, which are marked by a plurality of moral, religious, or philosophical doctrines —what he calls "comprehensive doctrines." Adherents of different comprehensive doctrines will endorse conflicting answers to a range of policy questions about which the state must nonetheless take a position. Central to Rawls's view is that the people who hold these different, often conflicting doctrines can nevertheless each be reasonable insofar as they acknowledge that others may hold different views and they are motivated to find and abide by fair terms of cooperation, provided that others will do the same.[5] Rawls contends that it is unjustifiable in this context of reasonable pluralism for the state to subject citizens

to coercive laws and policies based on comprehensive doctrines to which they do not adhere. Instead, political power must be exercised in ways that all citizens, as free and equal, could reasonably be expected to endorse.[6]

Rather than drawing on this or that comprehensive doctrine as a basis for the exercise of political power, Rawls writes, the state should ground its public policies in considerations drawn from society's shared political commitments. These commitments are reflected in what he terms the "public political culture," which includes "society's main institutions and the historical traditions of their interpretation."[7] The shared commitments may be discerned from sources such as, for example, important legal decisions or political documents. Citing the preamble to the US Constitution, Rawls mentions as examples of shared political commitments the values of "a more perfect union, justice, domestic tranquillity, the common defense, the general welfare, and the blessings of liberty for ourselves and our posterity."[8] Other examples of shared political commitments in a liberal democracy include the conception of citizens as free and equal and of society as a fair system of cooperation,[9] as well as ideas such as "liberty of conscience, or equal political liberties, or basic civil rights,"[10] and "public deliberation."[11] These sorts of ideas are appropriate to draw on in justifying the exercise of power in a liberal democracy because they are commitments we could expect any reasonable member of a liberal democracy to endorse.

An implication of this account of the legitimate exercise of state power, Rawls contends, is that citizens in positions to exercise or facilitate the exercise of political power—as, for example, legislators, judges, political candidates, and voters—have a duty of civility to justify answers to fundamental political questions by appeal to these shared political values, rather than their own comprehensive doctrines. He calls this the idea of "public reason," and he explains that it "provides a publicly recognized point of view from which all citizens can examine before one another whether their political and social institutions are just."[12]

Flanders draws on Rawls's account of public reason as the basis for his objection to retributivist justifications of punishment. Retributivist accounts vary in some significant ways, but what unites these theories is the idea that punishment is justified because it is an intrinsically appropriate, because deserved, response to criminal wrongdoing. On some accounts, punishment metes out deserved suffering; on others it communicates deserved censure; on still others, it constitutes the deserved removal of an unfair advantage the offender gained from his offense.[13] On any of these varieties of retributivism, the desert claim is central. Flanders contends that retributivism is inconsistent with public reason because the desert claim on which it relies is a claim about which citizens with different comprehensive doctrines may reasonably disagree. He writes:

I do not need to show that retribution is false (say, that there is no such thing as 'desert,' or that if there is desert, it does not justify the state's imposing hard treatment against people even if they deserve it). I need to show only that retribution is the subject of reasonable disagreement and so cannot be justified in terms that all can see as public reasons; that is, reasons for public policy.[14]

Flanders first considers what he calls the "bare intuition of retribution," the idea that wrongdoers deserve to suffer. Appeal to this intuition is central, for example, to Michael Moore's defense of retributivism: Moore recounts various heinous crimes and asks readers to consult their intuitions about whether it would be good for the perpetrators to be punished even if the punishment served no further purpose (such as crime reduction). Moore believes that most readers will share his intuition that punishment in such cases is intrinsically good, because deserved.[15] As Flanders points out, however, this idea is controversial, and an appeal to intuition "doesn't have much of an argument to offer against those who happen not to share the intuition."[16] Elsewhere Flanders writes:

It would be foolish to deny that there may be ideas of religious desert, or some such thing, so that the wicked deserve to be in hell and the saintly deserve to be in heaven. But such a conception would surely be ill-suited for politics.[17]

We can add to Flanders's critique the common objection that the notion of desert central to retributivism offers no plausible, nonarbitrary account of what an offender deserves in particular cases: that is, retributivism provides little or no actual sentencing guidance.[18] What's more, Flanders contends that the very notion of desert presupposes "rather controversial metaphysical ideas about freedom of the will."[19] Many people will disagree with the idea that we are free in the sense of being capable of choosing to do otherwise than we actually do.[20] For these people, the idea that we deserve to suffer for doing what we could not have avoided doing may seem objectionable.

Flanders also discusses the fairness-based variety of retributivism, according to which an offender, in committing a crime, gains an unfair advantage over other, law-abiding members of society, and punishment constitutes the deserved removal of this unfair advantage. As Flanders points out, this brand of retributivism is also controversial. Critics have objected, for example, to the ideas that an offender gains some advantage by committing a crime and that, whatever the advantage is, punishment somehow removes it.[21] Similarly, he critiques expressive retributivist accounts offered by Jean Hampton and others, on which punishment negates the offender's false claim of superiority over his victim that he expresses through his crime. Again drawing on familiar objections to this view, Flanders contends that people could reasonably reject such an account. He writes:

I think we may wonder whether punishment ever really is a good means to securing equality: that some people have to suffer in order to affirm the equal dignity of those who have been injured. Such an idea seems to me a clear source of reasonable disagreement.[22]

Flanders concludes that various prominent retributivist accounts are subject to reasonable disagreement, and thus retributivism cannot properly be the basis of exercises of coercive state power.

III. What Does Public Reason Require?

My aim in this section and the next is not primarily to defend retributivism as a justification of punishment. Rather, I contend that appeal to

Rawlsian public reason does not provide the resources to rule out retributivism as a candidate justification of punishment, or more generally, to settle questions of whether and why punishment is justified.

How should we understand the charge that retributivism violates public reason? On one commonly endorsed interpretation, public reason requires that exercises of political power not be based on reasons that people could reasonably reject, or about which there is reasonable disagreement. Flanders draws on this interpretation repeatedly, such as when he writes that retributivism is "a fairly controversial idea and relies on premises that other people in a liberal society could reject, and reasonably so," and that retributivism "is the subject of reasonable disagreement."[23] Retributivism is undeniably controversial. It is not unreasonable to reject it—indeed, many theorists have rejected it based on reasonable concerns such as those discussed earlier. But if public reason really requires that we not ground punishment policy in any reasons that are controversial, or that could be subject to reasonable disagreement, then we would arguably have no basis for justifying punishment at all. Every prominent justificatory account of punishment —be it based on retribution, deterrence, societal self-defense, offender reform or education, or other considerations, or some hybridization of these—has been controversial. Each has been the subject of reasonable disagreement. This is not to say that objections leveled against the various accounts are always decisive, but they are typically reasonable, in the sense that they are themselves based on considerations from our

public political culture—considerations that others, even those who adhere to different comprehensive doctrines, could be expected to endorse.

Flanders anticipates this objection that a prohibition on basing public policies on considerations about which there is reasonable disagreement will leave no grounds at all to justify punishment. He contends the rationale of specific deterrence— essentially, that punishing criminals will provide them a strong disincentive from offending the next time an opportunity presents itself—does not violate public reason. He writes:

The idea that we need to punish crime as a means of deterring others from committing more crimes—and specifically to prevent the one who has committed a crime from committing more crimes—[is] broadly acceptable, no matter a person's underlying theory of why we punish. It seems reasonable that a state would need mechanisms for societal protection and defense, of which punishment might be one. Even retributivists would presumably be troubled if a system of punishment based on desert led to more crime rather than less.[24]

Flanders adds that objections commonly leveled against deterrence views—namely, that they cannot in principle rule out punishment of the innocent or disproportionate punishment of the guilty—can be addressed simply by incorporating prohibitions on punishment of the innocent or disproportionate punishment of the guilty as side constraints in an account of punishment in which specific deterrence is the rationale. He contends that both of these constraints are "broadly acceptable" even to those who reject retributivism and the notion of desert.[25] In the end, then, Flanders concludes that a system of punishment aimed at

specific deterrence with constraints prohibiting punishment of the innocent or disproportionate punishment of the guilty cannot reasonably be rejected, and thus is consistent with public reason.

Of course, retributivists do reject deterrence, even specific deterrence, as a rationale for punishment, and they reject it even if punishment includes side-constraints prohibiting punishment of the innocent or disproportionate punishment of the guilty. Retributivists' objections are based on considerations such as equality, respect for persons, and moral rights, values that we generally take to be among the shared commitments in a liberal democracy. Consider, for example, Jeffrie Murphy's argument against utilitarian theories of punishment generally. On utilitarian theories, he contends, even a guilty person is punished "because of the instrumental value the action of punishment will have in the future. He is being used as a means to some future good."[26] Murphy argues that what a utilitarian theory of punishment cannot capture is "the notion of persons having rights":

Even if punishment of a person would have good consequences, what gives us (i.e., society) the moral right to inflict it? … What does this right to punish tell us about the status of the person to be punished—e.g., how are we to analyze his rights?[27]

Another objection to deterrence as a rationale for punishment, articulated by Antony Duff, contends that a deterrent system of punishment communicates with those subject to it in the language of prudential reasons, but that such reasons are "not appropriate for a liberal political community. The law of that

community, as its common law, must address its members in terms of the values it embodies—values to which they should, as members of the community, already be committed."[28] On Duff's account, a penal system that aims to secure obedience to the law through threats fails to treat citizens as autonomous moral agents. Similarly, Hegel likened deterrent punishment to shaking a stick at a dog as a threat; it treats a person "like a dog instead of with the freedom and respect due to him as a man."[29]

My aim here is not to defend these objections to deterrence as an aim of punishment. I have argued elsewhere that there are plausible responses to some of them.[30] But the objections are reasonable, in Rawls's sense. They are grounded in considerations—moral rights, equality, respect for citizens as autonomous moral agents—that are arguably among the shared political commitments of liberal democracies. Thus, they are objections that other members of the polity (at least insofar as they are themselves reasonable) could be expected to endorse irrespective of their own comprehensive doctrines. We can add to this that there is reasonable disagreement about whether, or to what degree, punishment is effective as a deterrent; thus, even those who accept that deterrence of potential offenders could be a permissible aim may reasonably disagree about punishment's efficacy as a means to that end.

The point, then, is that one could reasonably reject deterrence as a rationale for punishment. I believe the same can be said (although I do not spell it out here) about other candidate accounts of punishment,

based on self-defense, moral education, and so on. It might be tempting to think that this conclusion bodes well for abolitionists about punishment: If each candidate justification of punishment is subject to reasonable disagreement, and if being subject to reasonable disagreement disqualifies a view from serving as a proper ground of public policy, then perhaps punishment cannot be justified. But the case for abolitionism about punishment is itself the subject of reasonable disagreement. Thus, if we disqualify as grounds for important public policy choices any considerations that are subject to reasonable disagreement, then it appears that we will not have a legitimate basis on which to justify punishment or its abolition. (One might respond that justification is only needed for exercises of state power, not for non-exercises of it. I think, though, that given the enormous effect the abolition of punishment would likely have on citizens' lives—including potential victims of crime and their loved ones—the state's choice to abolish punishment would arguably need to be justifiable to citizens no less than its choice to impose punishment.)

The interpretation of public reason as requiring that the reasons we offer for public policy choices not be ones that could reasonably be rejected, or requiring that they are not subject to reasonable disagreement, seems to leave no admissible grounds for the imposition of punishment policies (or arguably for the abolition of them). This problem extends beyond penal theory. One of the key developments in Rawls's later work is that he acknowledges that his preferred political conception of justice, justice as fairness, is just one among many reasonable conceptions.[31] These various conceptions will be grounded in the ideas of citizens as free and equal and of society as a fair system of cooperation:

Yet since these ideas can be interpreted in various ways, we get different formulations of the principles of justice and different contents of public reason. Political conceptions differ also in how they order, or balance, political principles and values even when they specify the same ones.[32]

Insofar as each of these different political conceptions could reasonably be endorsed, each could also reasonably be rejected by adherents of one of the other reasonable conceptions. That is, the various conceptions each will be subject to reasonable disagreement. The interpretation of public reason as requiring that reasons for public policy not be such that they could be reasonably rejected, or not be the subject of reasonable disagreement, suggests that it is inconsistent with public reason to draw on any of the various reasonable political conceptions of justice as grounds for public policies. In fact, some philosophers have argued (persuasively, in my view) that public reason itself can be reasonably rejected; thus, on the "reasonable rejection" interpretation of what public reason requires, it appears that public reason is self-defeating.[33] We might conclude from all of this that public reason rules out drawing not only on retributivism in policy debates but also on other considerations that are subject to reasonable disagreement, such as deterrence, moral education, societal self-defense, and so on; that it also rules out drawing on any of the various reasonable political conceptions of justice; and in fact that

public reason itself cannot be justified on its own terms. I suggest, though, that we should instead consider an alternative, less restrictive interpretation of public reason.

Rather than interpreting public reason's requirements in terms of reasonable rejection or disagreement, I believe we do better to interpret it as requiring us to offer reasons that we could reasonably expect our fellow citizens—as free and equal citizens who are, like us, committed to finding and abiding by fair terms of cooperation—to endorse. It is reasonable to expect others to endorse the reasons we offer as grounds for public policies when these reasons are drawn from our shared political commitments. It is unreasonable to expect others to endorse reasons that are not drawn from our shared political commitments but are, instead, drawn from one's own comprehensive doctrine.

Rawls tends to write in terms of reasonable expectation of endorsement rather than of reasonable disagreement or rejection. For example, his liberal principle of legitimacy states:

Our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason.[34]

Similarly, he sets out the idea of public reason this way:

A citizen engages in public reason, then, when he or she deliberates within a framework of what he or she sincerely regards as the most reasonable political conception of justice, a conception that expresses political values that others, as free and equal citizens might also reasonably be expected reasonably to endorse.[35]

It is perhaps worth noting that the relevant sense of expectation here is normative rather than predictive. It may be normatively reasonable to expect others to endorse reasons I offer even if I can reasonably expect, in the predictive sense, that they will not endorse these reasons. By contrast, I might offer reasons that I cannot reasonably expect (in the normative sense) others to endorse, because the reasons are not drawn from our shared political commitments, even if I can reasonably expect (in the predictive sense) that many others will endorse these reasons, perhaps because we adhere to roughly the same comprehensive doctrine.

To see why it matters whether we understand public reason's requirements in terms of reasonable rejection (or disagreement) or in terms of reasonable expectation of endorsement, consider two different senses in which people might reasonably reject the reasons on which I base my support for some policy. On the one hand, they might reject the values on which my reasons are based if I have grounded my views in religious, ethical, or philosophical doctrines that are not among the shared values in our public political culture. On the other hand, even if my views are grounded in our shared values, they might reject my interpretation of these values, or my relative weighting of the various values, or the way I resolve tensions among them, or the implications I draw from these values for specific policy decisions. Liberty, equality, rights, fairness, respect for persons, public safety: these are all shared political commitments in a liberal democracy, but citizens can reasonably disagree about how to apply

these values to policy questions. In either of these ways, someone could reasonably reject the reasons I offer for some policy position. The rejections could be reasonable insofar as they are made in the spirit of desiring to find fair terms of social cooperation among free and equal citizens, and thus are themselves grounded in shared values in our public political culture.

Even though people could reasonably reject my reasons in either of the ways just mentioned, the second type of rejection—based on differing inferences about how the values apply to the policy question—is consistent with my maintaining a reasonable expectation of others that they endorse my reasons. When I draw on our shared political commitments as the basis of my reasons in support of some policy, it is reasonable for me to expect others to endorse these reasons even if it turns out, because they draw different inferences about how the shared values should apply in this case, that they reasonably reject my reasons. There is nothing inconsistent, then, with holding both that I can reasonably expect others to endorse my reasons in favor of some policy position and also that they can reasonably reject these reasons; this is just to acknowledge that there can be reasonable disagreement within the context of public reason itself. The requirement to offer reasons that we could reasonably expect others to endorse is therefore more permissive than the requirement to offer reasons that could not reasonably be rejected.

Importantly, the sort of endorsement that I can expect of others (if I have grounded my reasons in shared political commitments) is not agreement that my preferred policy decision or my reasons for endorsing it are correct, or true. Because reasonable people can interpret our shared political commitments in various, reasonable ways, it would be unreasonable for me to expect others to agree with my particular interpretation of these shared commitments, or to agree that my conclusions about the commitments' implications for particular policy issues are true. But insofar as I ground my principles in our shared political values, I thus demonstrate a commitment to seek and abide by fair terms of cooperation that are justifiable to others as free and equal members of the polity. It is reasonable, in such circumstances, for me to expect others to demonstrate a similar commitment to seek and abide by fair terms of cooperation. Thus, it is reasonable for me to expect others to endorse my reasons in the sense of accepting that, insofar as they are derived from our shared political commitments, these reasons are legitimate as a basis for public policy.[36]

By contrast, when I draw on my own religious, ethical, or philosophical commitments that are not among our shared political values, I fail to demonstrate a desire to seek and abide by fair terms of cooperation that are justifiable to others. It is then unreasonable for me to expect others to accept such reasons as legitimate grounds for public policy. If our reasons for supporting some policy are grounded in some religious doctrine, for example, it would be unreasonable to expect non-adherents to that religion to endorse these reasons. Thus, although the "reasonable expectation of endorsement" standard is more permissive than the "reasonable rejection" standard, it is not entirely permissive.

Returning to the question of punishment's justification, then, the question we should ask is not whether retributivism, deterrence theories, or other accounts are subject to reasonable disagreement, or could reasonably be rejected, but instead whether they can be derived from the shared political commitments of a liberal democracy. I think it should be fairly uncontroversial that deterrence justifications of punishment, grounded in considerations such as public safety and rights protections, meet this condition. This is not to deny that there can be reasonable disagreement about whether deterrence accounts properly realize these shared political commitments, or whether they properly weight these relative to other shared political commitments. But deterrence accounts, I think we should acknowledge, are derived from considerations that are among the shared political commitments of a liberal democracy, and thus it would not be unreasonable to expect citizens to endorse such accounts.

But what about retributivist accounts? Are these accounts actually grounded in elements of comprehensive doctrines that cannot be found among our shared political commitments, or can they be derived from our shared commitments? In the next section, I contend that at least some prominent retributivist accounts can be derived from our shared political commitments, and thus, despite being subject to reasonable disagreement, are consistent with public reason.

IV. Retributivism from Our Shared Political Commitments

I wrote earlier that what unifies the fairly broad range of retributivist theories is the notion that punishment is a deserved response to crime. The relevant question for present purposes, then, is whether this notion of a person's deserving punishment can be derived from our shared political commitments. Determining this requires that we take a closer look at what the retributivist desert claim is actually saying.

If we say that someone deserves some treatment, this is to say, first, that the treatment is in some sense appropriate. Joel Feinberg writes that person's desert of X is always a *pro tanto* reason for giving him X.[37] John Kleinig characterizes the desert claim as expressing, depending on the context, that the person "ought to get or suffer" X, or that "it would be a good thing" to give the person X.[38] A desert claim says more than this, however. It says that treating the person in a given way is not merely appropriate but *intrinsically* appropriate. The person ought to get X, or it would be a good thing to give the person X, not because of the valuable consequences this would produce, but rather "*in virtue of* characteristics or acts" of the person.[39] Similarly, Margaret Falls writes that a person deserves X "based upon what already is or has been true of P's characteristics, abilities, and/or acts."[40] Furthermore, as Falls points out, retributivism is grounded in a particular form of desert claim, what she calls a claim about earned desert. She contrasts earned desert—essentially, what we come to deserve in virtue of "performing human and social functions well or poorly"[41]—with unearned

desert—what we deserve simply in virtue of our status as moral persons. Falls acknowledges that her notion of unearned desert may be controversial. But what is relevant for our purposes is that the central retributivist claim is an *earned* desert claim: an offender deserves to be punished in virtue of having committed an offense.

We can thus cash out the central retributivist desert claim as saying that punishment constitutes intrinsically appropriate treatment of an offender in virtue of having committed an offense. What is it about punishment, though, that is intrinsically appropriate in this way? Some retributivist accounts focus on suffering: it is intrinsically appropriate that offenders are made to suffer. Such accounts typically use thought experiments to pump intuitions about the intrinsic value of deserved suffering. Michael Moore's account, mentioned earlier, is perhaps the most prominent version of deserved-suffering retributivism.[42] Moore writes, for example:

> [I]magine an offender who does a serious wrong in a very culpable way—e.g., Dostoevsky's Russian nobleman in *The Brothers Karamazov,* who turns loose his dogs to tear apart a young boy before the eyes of the boy's mother; imagine further that circumstances are such … that no non-retributive purpose would be served by punishing this offender. Now imagine two variations: (1) you are that offender; (2) someone else is that offender. Question: should you or the other offender be punished, even though no other social good will thereby be achieved? The retributivist's 'yes' runs deep for most people.[43]

The intuition on which Moore draws here — that there is intrinsic value in the suffering of the wicked — will indeed be powerful for many people. The question, however, from the perspective of public reason, is whether this intuition can be found among the shared commitments of a liberal political community. I am skeptical about this, but I do not pursue this point here. I simply note that if a proponent of such a view were interested in demonstrating that the view is consistent with public reason, they would need to establish that this intuition about the intrinsic value of deserved suffering can be found among the shared commitments in a liberal democratic polity. Rather than focus on this version of retributivism, however, I turn instead to other retributivist accounts that I think more clearly can be grounded in the shared commitments of a liberal democracy. Specifically, I have in mind versions of retributivism on which punishment is justifiable as the deserved removal of an unfair advantage, as a deserved communication of censure, or as a repudiation of an offender's representation of superiority to his victim.

Consider, first, the fair play version of retributivism. This account begins with a conception of society as a cooperative enterprise, where if we each accept some burdens associated with constraining our behavior to comply with the rules, we will all benefit. One feature of larger cooperative enterprises, however, is that a small number of people can opt out of contributing without having much (if any) effect on the benefits generated by the enterprise, to themselves as well as others. In other words, free riding is possible. Some theorists have argued that in such circumstances, when one person accepts

benefits made possible by others' participation in the enterprise, that person has an obligation to reciprocate by doing her fair share. The fair play account was first offered as theory of political obligation, but some scholars have extended the account in an attempt to justify punishment.[44] The extension runs this way: When a person commits a crime, she gains an unfair advantage over other members of the political community. She benefits, as others have, from general compliance with the laws. But she further benefits by not complying with the law when it suits her not to do so. On the fair play account, punishment is intrinsically appropriate as a way of removing the unfair advantage the offender gains from her crime.

As mentioned earlier, the fair play theory has faced a number of criticisms, but for present purposes we need not examine these in detail. We are not asking, after all, whether the fair play theory is correct; we are asking only if it can be derived from our shared political commitments. I believe it can. Rawls himself tells us that in a liberal democracy, one of the basic ideas implicit in the public political culture is that of society as a "fair system of cooperation between free and equal persons."[45] From this shared commitment, the fair play theorist argues that it is unfair for some to free ride by accepting benefits made possible by others without also shouldering their share of the burdens. Then the theorist adds that when some do unfairly gain benefits at the expense of others, it is intrinsically appropriate that society should take steps to remove those unfair benefits, and that punishment is an appropriate

way to do this. In my view, the fair play theory is the sort of account that we could reasonably expect other members of the polity, similarly committed to the idea of society as a fair system of cooperation, to endorse.

Many will not endorse the fair play theory, of course (there will be reasonable disagreement about it), because they will draw different conclusions about what the idea of society as a fair system of cooperation can tell us about state punishment. But reasonable disagreement about the policy implications of the idea of society as a fair system of cooperation is consistent with a reasonable expectation by proponents of the fair play theory that other citizens endorse it. The theory does not depend, as far as I can tell, on religious, ethical, or philosophical doctrines that it would be unreasonable to expect others to endorse. It simply draws on a commitment to fairness, a belief that offenders gain an unfair advantage over everyone else, and a contention that punishment is an appropriate way of restoring a fair balance in such cases. These latter two claims about the implications of a commitment to fairness are controversial, but I have argued that reasonable disagreement about the correct derivation of public policy positions from our shared political commitments is consistent with the idea of public reason. Put simply, although we can debate whether fair play retributivism succeeds as a justification of punishment, it is not unreasonable in the Rawlsian sense.

Another prominent version of retributivism, developed most notably by Antony Duff, views punishment as an intrinsically appropriate communication of censure in

response to public wrongs. This account begins with the notion of the criminal law as a communicative enterprise, and a concern to treat members of the polity as autonomous moral agents rather than merely as things to be controlled or molded. Treating citizens as autonomous agents requires that the law attempt not merely to persuade them to refrain from criminal conduct, but to persuade them to refrain from it because it is wrong. And when people do commit crimes, taking these violations seriously involves responding to them with censure. As Duff writes, not to censure public wrongs would be "to undermine—by implication to go back on —its declaration that such conduct is wrong."[46] Further, he writes, censure of conduct declared to be wrong "is owed to its victims, as manifesting concern for them and for their wronged condition that the declaration itself expressed."[47]

Having argued that state censure of public wrongs is intrinsically appropriate, Duff then maintains that part of what it is to censure someone is to urge the person to repent, reform, and reconcile with those he has wronged and the community generally. These are not further aims contingently linked to censure but rather are aims internal to censure. The hard treatment of punishment serves the aims of repentance and reform by helping to focus the offender's attention on the wrong, and it serves the aim of reconciliation by acting as an opportunity for the offender to make amends for the wrongdoing.[48]

In another retributivist account that is in some respects similar to Duff's, Margaret Falls argues that part of what respect for moral persons requires is that we hold people accountable for their behavior. Like Duff, she contends that when people are guilty of crimes, it is intrinsically appropriate to hold them accountable by blaming them. But in communicating this blame, or censure, sometimes words are not enough. Falls writes, "Just as calmly telling a friend she ought not to have lied to us communicates neither the pain she has caused nor our unqualified insistence that we not be so treated, so the state's verbal or written reprimand with attached explanation would be inadequate."[49] Thus, sometimes burdensome treatment is necessary to convey appropriate measures of blame.

Here again, I think even those who disagree with Duff's and Falls's accounts should agree that they are reasonable in Rawls's sense. They both begin with a concern to respect fellow members of the community as autonomous moral agents. In a liberal democratic society, which as Rawls says is implicitly committed to a conception of citizens as free and equal, respect for fellow community members as autonomous moral agents seems readily evident among the shared political commitments. Both Duff's and Falls's accounts also depend on the idea that treating others as autonomous moral agents can mean, among other things, blaming them when they are guilty of wrongdoing. As discussed, they then develop this notion in different ways to arrive at the conclusion that state punishment can be an intrinsically appropriate response to public wrongs. Now, can we expect that others will reasonably reject these accounts? Of course. But would it be unreasonable for Duff or Falls to

expect other members of a community committed to treating each other with respect as autonomous moral agents to endorse their respective accounts? I do not think so. As far as I can tell, there is nothing in either account that is unacceptable to other members of a liberal democratic community in the way that, for example, Catholic doctrine would be unacceptable to Hindu members of the community as a reason for coercive state policies.

As another example, consider Hampton's retributivist account. Hampton contends that punishment is intrinsically justified as a way of affirming a victim's equal dignity in response to conduct that causes moral injury, which is to say that the conduct "expresses and does damage to the acknowledgment and realization of the value of the victim."[50] As mentioned before, Flanders objects to the idea that "punishment ever really is a good means to securing equality: that some people have to suffer in order to affirm the equal dignity of those who have been injured."[51] In my view, Flanders's disagreement with Hampton is best interpreted as reasonable disagreement about how to interpret the values of equality and dignity, and what implications these values have for penal policy. But this disagreement is consistent with acknowledging that equality and dignity are the right sorts of notions to draw on in policy debates, as

they are among our shared political commitments. Likewise, Flanders points out that there is reasonable disagreement about whether or in what sense people have the sort of free will on which notions of desert seem to rely.[52] Nevertheless, a central commitment in a liberal democracy is to the conception of citizens as free and equal. It is thus not unreasonable to expect others to endorse as legitimate our preferred account of what freedom or equality imply for penal policy, as these values are among our shared political commitments. There will be reasonable disagreement about how to interpret freedom and equality, how to weight them relative to each other, or how to reconcile tensions between them, but these disagreements are all consistent with public reason.

Ultimately, then, I think at least some prominent versions of retributivism are consistent with public reason. They are accounts that people could reasonably reject, insofar as they rely on contestable claims about how to interpret and apply values such as dignity, fairness, respect for persons, equality, and so on. But insofar as they are grounded in these sorts of values, which are among our shared political commitments, they are also accounts that their adherents could reasonably expect other members of their political community to endorse.

V. Objections

One might object to my account that although Rawls did not write extensively about punishment, we do have some evidence for his views

about the justification of punishment generally, and about retributivism in particular, and this evidence does not support the idea that

retributivism is consistent with a Rawlsian account. The first piece of evidence comes from Rawls's most focused discussion of the justification of punishment, in his 1955 article "Two Concepts of Rules." The account he develops in this paper is often cited as a hybrid view of punishment, but in my view it is best interpreted as a version of rule utilitarianism, according to which utilitarian considerations justify implementing a system of punishment that incorporates retributivist prohibitions on punishment of the innocent or disproportionate punishment of the guilty. On Rawls's account, the rule (the institution of punishment incorporating the retributivist constraints) is justified insofar as it generates more overall net utility than available alternatives, and the individual actions (particular impositions of punishment governed by the retributivist constraints) are justified insofar as they conform to the rule. Retributivism clearly is relegated to a secondary role on this account, as the retributivist constraints are justified only by their conduciveness to the institution's utilitarian rationale.

The second piece of evidence comes from *A Theory of Justice*, where Rawls seems to reject the notion of pre-institutional desert. He writes that

no one deserves his place in the distribution of native endowments, any more than one deserves one's initial starting place in society. The assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit.[53]

Retributivist theories are grounded on the notion that punishment is a deserved response to wrongdoing, and the notion of desert at issue is apparently pre-institutional, or else appeal to desert as a justification of penal institutions would be question-begging. Thus, it appears that Rawls, in rejecting the idea of pre-institutional desert, is thereby rejecting retributivism as well.

There are at least two ways of responding to this first line of objection. The first response would be to question whether the two pieces of evidence mentioned really show that Rawls's considered views about punishment were opposed to retributivism. Considering first his account in "Two Concepts of Rules," we might wonder whether his views about punishment changed from this early rule utilitarian account, especially given the powerful arguments he levels against utilitarianism in *A Theory of Justice*.[54] Regarding his apparent rejection of pre-institutional desert in *A Theory of Justice*, Rawls's actual views about pre-institutional desert are the subject of significant philosophical debate. Some scholars deny that Rawls rejects the notion of pre-institutional desert[55]; if this is correct, then his account might not be inconsistent with retributivism. Thus, with respect to either his early discussion, or punishment, or his discussion of pre-institutional desert, we might deny that these actually indicate that Rawls's considered views involved a rejection of retributivism.

I mention the first line of response only briefly, because I believe a more promising second line of response is available: Regardless of whether Rawls's own preferred political conception of justice, "justice as fairness," rejects retributivism, he makes clear in his later writing, as

we discussed earlier, that this conception is only one among many that may be consistent with public reason:

[T]he content of public reason is given by a family of political conceptions of justice, and not by a single one. There are many liberalisms and related views, and therefore many forms of public reasons specified by a family of reasonable political conceptions of justice. Of these, justice as fairness, whatever its merits, is but one.[56]

Thus, even if Rawls himself believed that retributivism fails as a justification of punishment, the relevant question is whether an account of retributivism can be developed from the shared commitments of a liberal political community. I have suggested that at least some versions of retributivism are consistent with public reason in this regard.

A second line of objection, however, holds that it is not enough that a theory be grounded in shared political commitments. To be consistent with public reason, the interpretation of these commitments and the lines of reasoning from the commitments to their implications for policy questions must also themselves be shared public commitments. So even if fair play theorist's conception of society as a fair system of cooperation between free and equal persons is among our shared political commitments, the line of reasoning from this starting point to the justification of punishment—namely, that offenders gain an unfair advantage relative to other community members and that punishment is an appropriate way to remove this unfair advantage—is not itself among our shared commitments. Even if communicative accounts such as Duff's and Falls's

begin with a consideration that is among our shared political commitments—namely, a concern to respect those who commit crimes as fellow members of the polity, or as autonomous moral agents—these accounts then hold that treating offenders in such ways requires censuring them, and that punishment can be an appropriate way to convey this censure. These latter claims are not among our shared political commitments. And even if Hampton's starting point—a concern with affirming the equal dignity of those who are victims of crime—can be found among the shared commitments in a liberal democracy, her further claim that punishment is appropriate or effective in serving this function is not among our shared commitments. Perhaps, then, retributivist accounts are inconsistent with public reason not because their starting points are not among our shared political commitments, but because their lines of reasoning from those starting points to conclusions about penal policy do not consist entirely of shared commitments.

The problem with this line of objection is that, again, it sets too high a standard for what public reason requires. It would effectively rule out any normative theories of punishment, as any of these theories will offer certain interpretations of shared values, or draw certain inferences, that will not themselves be among our shared political commitments. For example, as discussed earlier, deterrence theories typically begin from a concern to protect public safety, which I believe is fairly clearly among the shared political commitments in a liberal democracy. But deterrence theories rely on a particular interpretation of what a

commitment to public safety implies —namely, that subjecting some people to intentionally onerous treatment to maintain a credible deterrent threat is an appropriate means of protecting public safety. This account of the implications of a commitment to public safety is not itself among our shared political commitments. It is a particular interpretation that is subject to reasonable disagreement even among those who share a commitment to the importance of ensuring public safety. I suggest that the same will be true of other theories of punishment, grounded in considerations of offender reform, societal self-defense, rights forfeiture, and so on (although space prevents me from analyzing each of these here). Any of these theories will involve lines of reasoning that are not themselves among our shared political commitments, and thus they will be subject to reasonable disagreement. Indeed, this is precisely the source of reasonable disagreement in the context of public reason. I maintain, therefore, that public reason requires only that our reasons for policy decisions be grounded in shared public commitments; it does not require, further, that our interpretations of these commitments or the inferences we draw from them to concrete policy prescriptions are themselves shared commitments.

A third possible line of objection might contend that whereas the "reasonable rejection" interpretation of public reason may rule out too much, the "reasonable expectation of endorsement" interpretation rules out too little. That is, if theories of punishment grounded in considerations as diverse as deterrence, fair play, deserved censure, societal self-defense and others are all consistent with public reason, then we might wonder if public reason does any actual heavy lifting in guiding policy debates. We should acknowledge, I believe, that public reason will accommodate a fairly robust range of conflicting views about what public policies and institutions are justified. Rawls himself seemed to endorse this view. In considering how much political philosophy can do to settle questions about what sort of political institutions are legitimate, he wrote "it is likely that the most that can be done is to narrow the range of public disagreement"[57] and that "the public political culture may be of two minds even at a very deep level."[58]

As I mentioned before, however, this is not to say that public reason is entirely permissive. Even on this more permissive interpretation, public reason rules out, for example, appeal to religious doctrine as a basis for public policy. It also rules out appeal to comprehensive philosophical doctrines, such as comprehensive utilitarian or Kantian moral theories. Rawls contends that these comprehensive doctrines, insofar as they are reasonable, may be consistent with various shared political values on which public policy could appropriately be grounded. But the comprehensive doctrines themselves are "not suitable to provide a public basis of justification."[59] Thus, even on the interpretation I endorse, public reason does exclude some types of reasons as grounds for public policy.

VI. Conclusion

I have argued that Rawls's notion of public reason, on its most plausible construal, does not provide a basis for ruling out retributivist reasons for punishment, at least on some fairly well-known retributivist accounts, insofar as they can be derived from our shared political commitments. I have also suggested that the idea of public reason offers fairly modest resources for settling debates in punishment theory more generally. Accounts grounded in considerations such as deterrence, moral education, societal self-defense, or right forfeiture; pluralist theories; or abolitionist accounts will all often be consistent with public reason, insofar as they are grounded in shared political commitments. This is not to say that all of these accounts are correct, or to deny that some accounts are more plausible than others. But there is reasonable disagreement about their comparative plausibility, and this reasonable disagreement can be accommodated within the context of public reason.

Notes

1 The notable exception is his discussion of punishment in Rawls, "Two Concepts of Rules"; I discuss this account in section V.

2 See, e.g., Dolovich, "Legitimate Punishment"; Brettschneider, "Rights of the Guilty"; Matravers, "Political Theory"; Chiao, *Criminal Law.*

3 Non-Rawlsian accounts of punishment are too numerous to cite here, but it is perhaps worth noting a few accounts that are grounded not in Rawlsian liberal theory but rather in communitarian political theories: Yankah, "Republican Responsibility"; Dagger, "Republicanism and the Foundations"; and Duff, *Punishment, Communication, and Community,* 35–73.

4 See Flanders, "Can Retributivism Be Saved?"; "Punishment, Liberalism"; "Political Philosophy and Punishment."

5 See Rawls, *Political Liberalism,* 49; "Public Reason Revisited," 770.

6 See Rawls, *Political Liberalism,* 137.

7 Rawls, "Justice as Fairness," 228.

8 Rawls, "Public Reason Revisited," 776.

9 See e.g., Rawls, "Justice as Fairness," 231–34; "Public Reason Revisited," 770.

10 Rawls, *Political Liberalism,* 155.

11 Rawls, "Public Reason Revisited," 772.

12 Rawls, *Political Liberalism,* 9.

13 See, respectively, Moore, *Placing Blame*; Duff, *Punishment, Communication, and Community*; Morris, "Persons and Punishment."

14 Flanders, "Punishment, Liberalism," 66.

15 See Moore, *Placing Blame,* 98–100.

16 Flanders, "Punishment, Liberalism," 67.

17 Flanders, "Can Retributivism Be Saved?" 309.

18 See, e.g., Shafer-Landau, "Retributivism and Desert."

19 Flanders, "Punishment, Liberalism," 67.

20 See, e.g., Caruso, *Rejecting Retributivism*; Pereboom, "Free Will Skepticism."

21 See Flanders, "Punishment, Liberalism," 67; see also, e.g., Falls, "Retribution, Reciprocity, and Respect," 30–35; Dolinko, "Some Thoughts about Retributivism," 545–49; Hoskins, "Fair Play," 56–58.

22 Flanders, "Punishment, Liberalism," 67; see also, e.g., Dolinko, "Some Thoughts about Retributivism," 549–54.

23 Flanders, "Punishment, Liberalism," 66; see also Flanders, "Political Philosophy and Punishment," 540–41.

24 Flanders, "Punishment, Liberalism," 69.

25 Ibid.

26 Murphy, "Marxism and Retribution," 219.

27 Ibid., 220.

28 Duff, *Punishment, Communication, and Community*, 78; see also Duff, *Trials and Punishments*, 178–86.

29 Hegel, *Philosophy of Right*, 246.

30 See Hoskins, "Deterrent Punishment."

31 See Rawls, "Public Reason Revisited," 773–74.

32 Ibid., 774.

33 See, e.g., Wall, "Is Public Justification Self-Defeating?"; Mang, "Public Reason"; but cf. Bespalov, "Against Public Reason's."

34 Rawls, *Political Liberalism*, 137.

35 Rawls, "Public Reason Revisited," 773.

36 See Bespalov, "Against Public Reason's," 621.

37 See Feinberg, "Justice and Personal Desert," 224.

38 Kleinig, "Concept of Desert," 76.

39 Ibid.

40 Falls, "Retribution, Reciprocity, and Respect," 38.

41 Ibid., 40.

42 See Moore, *Placing Blame*, 98–100. For similar accounts, see, e.g., Zaibert, *Rethinking Punishment*, chapter 2; Alexander and Ferzan, *Reflections on Crime*, 184–85.

43 Moore, "Justifying Retributivism," 25.

44 E.g., Morris, "Persons and Punishment"; Dagger, "Playing Fair with Punishment"; Dagger, "Punishment as Fair Play".

45 Rawls, "Justice as Fairness," 231.

46 Duff, *Punishment, Communication, and Community*, 28.

47 Ibid.

48 See ibid., 106–12.

49 Falls, "Retribution, Reciprocity, and Respect," 42–43.

50 Hampton, *Intrinsic Worth of Persons*, 134.

51 Flanders, "Punishment, Liberalism," 67.

52 See ibid.

53 Rawls, *A Theory of Justice*, 104.

54 See ibid., esp. 22–27.

55 See, e.g., Moriarty, "Desert and Distributive Justice"; but cf. Logar, "Rawls's Rejection."

56 Rawls, "Public Reason Revisited," 773–74.

57 Rawls, "Justice as Fairness," 228.

58 Ibid., 228–29.

59 Rawls, *Political Liberalism*, 99.

ORCID

Zachary Hoskins http://orcid.org/0000-0003-4445-1412

Bibliography

Alexander, Larry, and Kimberly Kessler Ferzan. *Reflections on Crime and Culpability: Problems and Puzzles*. Cambridge: Cambridge University Press, 2018.

Bespalov, Andrei. "Against Public Reason's Alleged Self-Defeat." *Law and Philosophy* 40 (2021): 617–644. doi:10.1007/s10982-021-09418-6.

Brettschneider, Corey. "The Rights of the Guilty: Punishment and Political Legitimacy." *Political Theory* 35 (2007): 175–199. doi:10.1177/0090591706297714.

Caruso, Gregg D. *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*. Cambridge: Cambridge University Press, 2021.

Chiao, Vincent. *Criminal Law in the Age of the Administrative State*. New York: Oxford University Press, 2019.

Dagger, Richard. "Playing Fair with Punishment." *Ethics* 103, no. 3 (1993): 473–488. doi:10.1086/293522.

Dagger, Richard. "Punishment as Fair Play." *Res Publica* 14 (2008): 259–275. doi:10.1007/s11158-008-9071-1.

Dagger, Richard. "Republicanism and the Foundations of Criminal Law." In *Philosophical Foundations of Criminal Law*, Edited by R. A. Duff and Stuart Green, 44–66. Oxford: Oxford University Press, 2011.

Dolinko, David. "Some Thoughts about Retributivism." *Ethics* 101, no. 3 (1991): 537–559. doi:10.1086/293316.

Dolovich, Sharon. "Legitimate Punishment in a Liberal Democracy." *Buffalo Criminal Law Review* 7 (2004): 307–442. doi:10.1525/nclr.2004.7.2.307.

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

Duff, R. A. *Trials and Punishments*. Cambridge: Cambridge University Press, 1986.

Falls, M. Margaret. "Retribution, Reciprocity, and Respect for Persons." *Law and Philosophy* 6, no. 1 (1987): 21–51. doi:10.1007/BF00142639.

Feinberg, Joel. "Justice and Personal Desert." In *Rights and Reasons: Essays in Honor of Carl Wellman*, Edited by Marilyn Friedman, Larry May, Kate Parsons, and Jennifer Stiff, 221–250. Dordrecht: Kluwer Academic Publishers, 2000.

Flanders, Chad. "Can Retributivism Be Saved?" *BYU Law Review* 2 (2014): 309–362.

Flanders, Chad. "Political Philosophy and Punishment." In The Palgrave Handbook of Applied Ethics and the Criminal Law, Edited by Larry Alexander and Kimberly Kessler Ferzan, 521–545. Cham, Switzerland: Palgrave Macmillan, 2019.

Flanders, Chad. "Punishment, Liberalism, and Public Reason." *Criminal Justice Ethics* 36, no. 1 (2017): 61–77. doi:10.1080/0731129X.2017.1302638.

Hampton, Jean. *The Intrinsic Worth of Persons: Contractarianism in Moral and Political Philosophy*. Edited by Daniel Farnham. Cambridge: Cambridge University Press, 2007.

Hegel, G. W. F. The Philosophy of Right. Translated by T. Knox. Oxford: Oxford University Press, 1942.

Hoskins, Zachary. "Deterrent Punishment and Respect for Persons." *Ohio State Journal of Criminal Law* 8, no. 2 (2011): 369–384.

Hoskins, Zachary. "Fair Play, Political Obligation, and Punishment." *Criminal Law and Philosophy* 5, no. 1 (2011): 53–71. doi:10.1007/s11572-010-9103-8.

Kleinig, John. "The Concept of Desert." *American Philosophical Quarterly* 8, no. 1 (1971): 71–78.

Logar, Tea. "Rawls's Rejection of Preinstitutional Desert." *Acta Analytica* 28 (2013): 483–494.

Mang, Franz. "Public Reason Can Be Reasonably Rejected." *Social Theory and Practice* 43, no. 2 (2017): 343–367. doi:10.5840/soctheorpract20172146.

Matravers, Matt. "Political Theory and the Criminal Law." In *Philosophical Foundations of Criminal Law*, Edited by R. A. Duff and Stuart Green, 67–82. Oxford: Oxford University Press, 2011.

Moore, Michael S. "Justifying Retributivism." *Israel Law Review* 27, 1&2 (1993): 15–49. doi: 10.1017/S0021223700016836.

Moore, Michael S. *Placing Blame: A Theory of Criminal Law*. Oxford: Oxford University Press, 1997.

Moriarty, Jeffrey. "Desert and Distributive Justice in *A Theory of Justice*." *Journal of Social Philosophy* 33, no. 1 (2002): 131–143.

Morris, Herbert. "Persons and Punishment." *The Monist* 52 (1968): 475–501.

Murphy, Jeffrie G. "Marxism and Retribution." *Philosophy & Public Affairs* 2, no. 3 (1973): 217–243.

Pereboom, Derk. "Free Will Skepticism and Criminal Punishment." In *The Future of Punishment*, Edited by T. A. Nadelhoffer, 49–78. Oxford: Oxford University Press, 2013.

Rawls, John. "The Idea of Public Reason Revisited." *The University of Chicago Law Review* 64, no. 3 (1997): 765–807. doi:10.2307/1600311.

Rawls, John. "Justice as Fairness: Political Not Metaphysical." *Philosophy & Public Affairs* 14, no. 3 (1985): 223–251.

Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Rawls, John. "Two Concepts of Rules." *The Philosophical Review* 64, no. 1 (1955): 3–32. doi:10.2307/2182230.

Shafer-Landau, Russ. "Retributivism and Desert." *Pacific Philosophical Quarterly* 81 (2000): 189–214. doi:10.1111/1468-0114.00102.

Wall, Steven. "Is Public Justification Self-Defeating?" *American Philosophical Quarterly* 39, no. 4 (2002): 385–394.

Yankah, Ekow N. "Republican responsibility in Criminal Law." *Criminal Law and Philosophy* 9 (2015): 457–475. doi:10.1007/s11572-013-9283-0.

Zaibert, Leo. *Rethinking Punishment*. New York: Cambridge University Press, 2018.