Crim Law and Philos
DOI 10.1007/s11572-013-9225-x

ORIGINAL PAPER

Facing the Consequences

Nathan Hanna

Abstract Many philosophers endorse deterrence justifications of legal punishment. According to these justifications, punishment is justified at least in part because it deters offenses. These justifications rely on empirical assumptions, e.g., that non-punitive enforcement can't deter or that it can't deter enough. I'll challenge these assumptions and argue that extant deterrence justifications of legal punishment fail. I begin by isolating, in a simplified form, important claims common to deterrence justifications. I then endorse an uncontroversial claim about punishment and explore its implications for enforcement. These implications undermine the simple versions of the deterrence claims. I then evaluate several modifications of the claims to see whether they can be improved upon. I argue that they can't easily be improved upon. In the process, I examine contemporary deterrence research and argue that it provides no support for deterrence justifications. I conclude by considering objections.

Keywords Punishment · Deterrence · Justification

Introduction

Consider what I'll call the Deterrence Thesis. Here's a preliminary statement of it.

D The fact that legal punishment deters is one reason why it is justified

N. Hanna (✉)
Department of English and Philosophy, Drexel University, MacAlister Hall Room 5044, Philadelphia, PA 19104, USA
e-mail: nhanna@drexel.edu

Published online: 06 April 2013

Many philosophers of punishment, including many non-consequentialists and many retributivists, think that a viable justification of punishment must appeal to such a claim.[1] Call such justifications *deterrence justifications*. These justifications rely on crucial empirical assumptions, e.g., that punishment deters substantially and that non-punitive enforcement can't. I'll challenge these justifications by challenging these assumptions.

Most philosophers think that these assumptions are obviously true. Some voice doubts about them, only to set the doubts aside as minor or beyond their expertise. A few press the doubts, but not always clearly or systematically. This is puzzling—and frustrating—given the popularity of deterrence justifications and the significance of these assumptions. Critics of deterrence justifications often associate them with consequentialist moral theories. They often focus on criticizing these theories instead of the assumptions. But not all deterrence justifications are obviously subject to the standard criticisms of these theories.

A better way to take on all deterrence justifications is to challenge their empirical assumptions. That's what I'll do. My aim is not to show that deterrence justifications can't succeed, just that extant versions of them don't and that philosophers of punishment should pay more attention to the empirical issues here. I also want to clarify what advocates of deterrence justifications have to show to vindicate their view.

I begin by isolating, in a simplified form, important claims common to deterrence justifications. I then endorse an uncontroversial claim about punishment and explore its implications for enforcement. These implications undermine the simple versions of the deterrence claims. I then evaluate and reject several modifications of the claims. In the process, I examine contemporary deterrence research and argue that it provides no support for deterrence justifications. I conclude by considering objections.

The Necessity Thesis

I'll give a brief overview of deterrence justifications and isolate some important claims they make.

Bentham endorses a utilitarian justification of punishment. He says that crime prevention "ought to be the chief end of punishment, as it is its real justification" (1843: 396). He insists that punishment is "in itself evil," that it is justified "only … in as far as it promises to exclude some greater evil," and that it is unjustified when it is "needless" (397). In other words, he thinks that if punishment didn't have these effects or if they could be secured in less harmful ways, punishment would be unjustified.

Utilitarian justifications of punishment have been extensively criticized. I won't reiterate the criticisms. Contemporary deterrence justifications try to avoid them. The most well known attempts to do this are due to John Rawls (1955) and H.L.A. Hart (1959). Rawls tries to show how rule utilitarians might try to justify punishment. Hart offers a similar justification. These justifications have also been extensively criticized. Again, I won't reiterate the criticisms.[2] For my purposes, only the following is significant. Both justifications retain an important qualification. On these views, punishment is justified, not just because of its deterrent effects, but because punishment is *necessary* to achieve these effects. Call this the Necessity Thesis.

[1] Throughout, I'll use *punishment* to mean *legal punishment*. Non-consequentialists or retributivists who think that it's necessary to appeal to something like D include Berman (2008), Husak (1992), Husak (forthcoming), Rawls (1971: 240), and Tadros (2011). Also see Rawls (1955) and Hart (1959).

[2] For an overview of all these justifications and standard criticisms see Boonin (2008: 39–77).

N Legal punishment is necessary to deter.

I state N simply, to match the initial statement of D. I'll examine possible modifications of them later.

Deterrence justifications are diverse, but all endorse something like N. It's obvious why. If N is true, it lends D substantial support. If N is false, the fact that punishment deters isn't itself a reason to prefer punishment to other ways of deterring. Advocates of deterrence justifications deny that there are other ways of deterring—or at least other ways of deterring enough.

I won't discuss specific justifications in detail. I'll just illustrate the general reliance on claims like N and examine standard defenses of it. I'll criticize these claims, along with claims like D. This is a way to challenge all deterrence justifications without having to deal with the idiosyncrasies of particular views.

In *A Theory of Justice*, Rawls explicitly appeals to a claim like N.

> It is reasonable to assume that even in a well-ordered society the coercive powers of government are to some degree necessary for the stability of social cooperation. … [A] coercive sovereign is presumably always necessary, even though in a well-ordered society sanctions are not severe and may never need to be imposed. Rather, the existence of effective penal machinery serves as men's security to one another. (1971: 240; cf. 315, 576)

Rawls makes an interesting move here—a bad one, as I'll make clear. He moves from the claim that coercive enforcement is necessary to maintain social stability to the claim that punishment is necessary to do so. He concludes that deliberators in the original position would authorize punishment.[3]

David Hoekema follows Rawls, but explicitly considers what enforcement options there are (1980: 248–49). He says that rejecting punishment would leave the deliberators with these options: a formal system of persuasion and censure or no enforcement at all. He argues that, given the options, they would choose punishment "because the harm they are likely to suffer if no such institution exists poses an unacceptable cost" (252).

> [Offenders'] punishment will be justified … by the necessity of an institution of punishment which includes their being punished in order to bring about a tolerable life for all (254).

On this view, punishment is justified because punishment is necessary to maintain tolerable social conditions. Hoekema thinks that this can be demonstrated from the armchair—by listing the alternatives to punishment and pointing out their obvious inadequacy. Call this the Inadequacy of Alternatives Thesis.

IA Non-punitive enforcement cannot deter

Again, I state IA simply. I'll examine modifications later.

Claims like N form the backbone of deterrence justifications. Advocates of these justifications typically either assume such claims or argue for them with Hoekema-style defenses of claims like IA. C.S. Nino, for example, says that the fact that punishment "is a

[3] This is a simplification. The deliberators don't consider these issues in the original position. They do so after choosing principles of justice—in a constitutional convention where the veil of ignorance is partially lifted (196–99, 240).

necessary and effective means of protecting the community against greater harms" provides, in conjunction with other considerations, "a prima facie moral justification" for it (1983: 299). Warren Quinn says that "common sense urges … that some form of civic punishment is necessary for a decent social order" (1985: 327). Anthony Ellis describes the options available to potential victims as follows: doing nothing, rational persuasion, or threatening retaliation (2003: 344). He says that we're justified in threatening and carrying out retaliation—in punishing, as he takes it—because doing so is necessary to provide disincentives to violence (343). And Victor Tadros says that punishment is "necessary for the maintenance of civil society" (2011: 285; cf. 298–302).

Appeals to claims like N aren't always so explicit, though. Daniel Farrell (1985, 1990), Phillip Montague (1995, 2002), and Erin Kelly (2009) appeal to principles of distributive justice in defense of claims like D. Here's one of Farrell's principles, slightly simplified.

> When someone S knowingly and wrongfully brings it about that someone else E must choose either to harm S or to be harmed herself, justice allows E to harm S, so long as the harm inflicted on S is roughly proportional to the harm that would otherwise be inflicted on E. (Farrell 1990: 302–3; cf. Montague 1995: 42, 2002: 26, Kelly 2009: 448).

Farrell thinks that this principle justifies self-defense. He also thinks that it grounds a principle that justifies punishment (1990: 316). By victimizing us, he says, offenders publicize our vulnerability, thereby increasing the risk that others will victimize us. So we're faced with a choice, he says: punish them, thereby deterring further attacks, or accept the greater risks. Within limits, he thinks we're entitled to punish offenders to reduce these risks.

Farrell's justification has a claim like N built into it: we "must choose" either to punish or to be subject to a greater—presumably substantial—risk of being wrongfully harmed. He takes punishment to be necessary to avert or substantially reduce this risk. And he holds that we're entitled to do what's necessary, within limits, to avert or substantially reduce it.

To summarize, I've called attention to the following three claims: The Inadequacy of Alternative Thesis, The Necessity Thesis, and the Deterrence Thesis.

IA Non-punitive enforcement cannot deter
N Legal punishment is necessary to deter
D The fact that legal punishment deters is one reason why it is justified

The fact that a diverse array of deterrence justifications appeal to claims like N and IA emphasizes the crucial role they play in these justifications. In the next section, I endorse a plausible claim about punishment. I explore its implications for enforcement and argue that IA, N, and D are false. Then I consider and reject several modifications of the claims.

The Harm Thesis

I endorse the following claim. David Boonin (2008) calls it the Harm Thesis.

H The aim to inflict harm is essential to punishment.[4]

[4] The term *aim* is meant to encompass both intentions (individual and collective) as well as non-mental analogues of intentions that institutions can have.

H is intuitive and widely accepted. I argue for it elsewhere (Hanna 2008, 2009a, b). David Boonin (2008: 3–28), Victor Tadros (2011: 1), and Michael Zimmerman (2011: 7–10) endorse versions of H, as do countless others.[5] H is compatible with many other popular claims about punishment and about what might justify it.[6] Punishment's advocates, including advocates of deterrence justifications, have good reasons to accept H and no obvious reason to reject it outright. I'll largely assume it and highlight some implications.

Before getting to these, though, I must address a set of objections to H that I haven't considered in print before. Different theorists say different things about what punishment aims to inflict. Among the suggestions are: burdens, deprivations, restrictions, pain, and suffering. Proponents of these views might reject H on the grounds that these things aren't—or at least need not be—harms. Some theorists also claim that people can benefit on the whole by being punished, e.g., by being reformed or morally educated (e.g., Hampton 1984). These theorists might reject H on the grounds that it entails that punishment can't do this.[7]

These objections rely on a certain conception of harm. To address them, I must clarify what *harm* means in H. I hold that punishment aims to inflict what Ben Bradley calls *prima facie* harms (2009: 65–68). He distinguishes these from what he calls *all things considered* harms. All things considered harms are "harms that are bad for a person taking into account all their effects on" her (2009: 66). By contrast, he says, prima facie harms are harms that (i) are intrinsically bad for a person, or (ii) bring about something intrinsically bad for her, or (iii) prevent something intrinsically good for her.

This distinction undermines the objections. Burdens, deprivations, restrictions, pain, and suffering are clearly prima facie harms. Those who say that punishment aims to inflict these things can't reject H by saying that these things aren't or need not be harms. The objection mistakenly takes H to be a claim about all things considered harms (this, at least, is the only reading on which the objection seems initially plausible). Ditto for the benefit objection. H is consistent with the claim that people can benefit on the whole by being punished. H just entails that if punishment can do this (qua punishment), it would just do so by—and in spite of—inflicting prima facie harm.[8]

H implies that the imposition of burdens, deprivations, and so on counts as punishment only if these things are imposed to inflict prima facie harm. Now, one might want to deny this claim too. But that's a hard objection to pull off. I've argued against it elsewhere (Hanna 2008, 2009a, b). Here, I'll just flatly state why I think it fails. For one, it's unmotivated and counterintuitive. For another, it makes it hard to distinguish punishment from other practices that are clearly non-punitive. It also makes it hard to explain important

[5] See, e.g., Barnett (1977: 280), Berman (2008: 271), Duff (2001: xiv), Feinberg (1963 [1970]: 67), Golash (2005: 45, 77–78), Lucas (1968: 207), Nino (1983: 292), and Sayre-McCord (2001: 2–3). For more references see Boonin (2008: 13, n14). Note that some of these authors don't use the word *harm* in their characterizations of punishment. More on this below.

[6] These include the claim that punishment aims to express criticism and the claims that it aims to impose burdens, deprivations, restrictions, pain, or suffering (more on this below). Note also that H says nothing about whether or why punishment is justified. Considerations of desert or deterrence, for example, may justify punishment, so far as H goes. H is compatible with Bentham's view that harm is an evil that we're justified in inflicting only if doing so is necessary to secure an outweighing good (*aim* as I'm using it doesn't mean *end* or *ultimate goal*). And it's compatible with the retributivist view that deserved harm is worth inflicting for its own sake.

[7] Thanks to an anonymous referee for pressing me to consider these objections.

[8] I say "in spite of" because prima facie harms will be all-things-considered harms unless they have good effects for the person that outweigh their prima facie harmfulness.

differences between punishment and these practices. Those interested in the details should look to my earlier work.

If you still think H is false, we can distinguish between *types* of punishment on the basis of their aims. If you deny H, take my argument in these terms, i.e., as challenging deterrence justifications of a certain type of punishment: punishment that aims to inflict prima facie harm. That's at least partly what deterrence justifications are meant to justify. And it's certainly something that contemporary enforcement systems do, partly for reasons of deterrence. So modifying my argument in this way won't change its substance. Either way, it poses a challenge to deterrence justifications of a certain way of conducting enforcement.

Now we can move on to H's implications for deterrence justifications. Here's the crucial implication I'll focus on: if H is true, there are more non-punitive options than philosophers typically assume. Standard enforcement techniques can be used without the aim to harm. These include but are not limited to probation, community service, compensation orders, fines, and forms of incapacitation like restraining orders, surveillance, and confinement. These techniques can be used to punish, but they don't have to be used that way. They aren't essentially punitive. H implies that any technique that doesn't require the aim to harm can be used non-punitively.

This undermines the standard argument for IA. Recall, the argument rejects non-punitive enforcement as obviously unable to deter. It assumes that such enforcement is limited to techniques like persuasion and criticism. It asserts that such techniques can't deter (or can't deter enough). For argument's sake, we can grant the claim about these techniques' deterrent capabilities. But these aren't the only non-punitive techniques.[9]

The standard argument for IA misrepresents the non-punitive options. One can endorse using standard techniques to pursue other aims while rejecting the aim to harm.[10] These techniques can and typically are used to do many things besides harming. Rejecting that aim doesn't preclude a rationale for their use. As we'll see, it just affects whether, when, and how they'll be used.

I won't defend a specific non-punitive enforcement proposal or a concrete set of aims.[11] Others do so, but it isn't necessary for my purposes (cf. Barnett 1977, Boonin 2008, Golash 2005, Sayre-McCord 2001).[12] I just want to challenge extant deterrence justifications, not to argue that no such justification can work. Extant justifications rely on assumptions about the comparative deterrent effectiveness of punishment and non-punitive enforcement. Defending these assumptions requires exploring potential non-punitive enforcement

[9] I don't know why philosophers often take criticism to be a paradigmatic non-punitive technique. Like the other techniques, verbal criticism can be punitive, but needn't be. Maybe these philosophers are mistakenly assuming that punishment is inherently violent or physically coercive. Such an assumption implies that the punitive options are more limited than they really are.

[10] And we shouldn't forget the sorts of non-standard techniques endorsed by advocates of restorative justice. See, e.g., Braithwaite (2002).

[11] Candidate aims include incapacitating the dangerous, securing compensation, expressing criticism, reforming offenders, and reconciling offenders with victims and other people. I won't endorse a particular set of aims, but I think pursuit of such aims is justified, within limits. Contemporary enforcement systems pursue such aims and I know of no advocate of deterrence justifications who rejects them wholesale.

[12] Following Barnett (1977), Boonin (2008) considers abolishing punishment and replacing it with victim restitution. He says that standard techniques like incarceration can furnish restitution and that offenders owe restitution to various parties for most standard offenses. I don't endorse his proposal (technically, neither does he; he thinks restitution is more defensible than punishment and that this demonstrates that punishment is unjustified). But it's a good example of the *kind* of proposal I'm discussing.

proposals. Standard arguments for IA preclude this by drastically misrepresenting the enforcement options. One of my aims is to emphasize the need for further work here.

Returning to H's implications, we've seen that the standard argument for IA fails. This doesn't show that IA, N, and D are false. But given H's implications, it's easy to show this. Here's the argument. Rejecting the aim to harm doesn't commit one to harmless enforcement. Enforcement that doesn't aim to harm can still harm incidentally because pursuit of other aims can cause incidental harm. For example, confining someone harms her even if it's only meant to incapacitate her. The harm here is incidental to the aim being pursued. Granting the claim that harm deters, the use of imprisonment and other enforcement techniques can deter even if they aren't used punitively. So there's no reason to assume that non-punitive enforcement can't deter substantially, let alone at all. Incidental harm can deter and non-punitive enforcement can deter because it may inflict such harms in pursuit of its aims.[13]

This shows that IA, N, and D are false. The fact that punishment deters doesn't distinguish it from non-punitive enforcement. So that fact isn't a reason to think that punishment is justified. Advocates of deterrence justifications think that its deterrent power helps to distinguish it from alternatives. So these justifications need more specific claims about punishment's deterrent power and about the justificatory significance of that power.

I'll consider such claims after considering an objection. My reply suggests ways to modify IA, N, and D.

An Objection

One might object that non-punitive enforcement of the sort I've described isn't different from punishment in any morally significant way. It still harms. Punishment is problematic because it harms. So, the objection goes, I've given no reason to think that the sort of enforcement I've described is any less problematic. The objection insists that the aim to harm poses no special justificatory problem. If so, my argument against IA, N, and D either fails or has no significant implications for the punishment debate.

One way to respond would be to endorse the Doctrine of Double Effect. If the doctrine is true, the aim to harm is intrinsically morally significant. Punishment abolitionists often appeal to the doctrine.[14] I don't. This isn't because I reject it. I'm agnostic about it. Instead, I appeal to something else: the effect that the aim to harm has on enforcement.

The effect can be illustrated as follows.[15] Consider enforcement that rejects the aim to harm and that aims to minimize the incidental harm it inflicts in pursuit of its aims. Take a system of punishment and compare it with an otherwise identical system like this. By "otherwise identical" I mean that they share all other aims, have the same capabilities and resources, operate in the same conditions, etc. This comparison isolates the aim to harm so that its effects can be assessed. I claim that the non-punitive system might inflict less harm

[13] At least some of the deterrence would be incidental. But not all of it must be. Such a system can aim to deter. It just wouldn't pursue that aim by aiming to harm. It could, for example, try to deter by doing things to amplify incidental harm's deterrent effects. I'll come back to this point later.

[14] See Boonin (2008: 15–16, 28–29, 61–62), Golash (2005: 45–48), Sayre-McCord (2001: 507), and Zimmerman (2011: 159–165).

[15] The following draws from Hanna (2012).

on the whole and wouldn't inflict more harm on the whole. Note that I'm not saying that such a system *would* inflict less harm, just that it *might*.[16] This is an important difference.[17]

The differing aims can affect enforcement in various ways. They can, for example, affect which sentences are imposed and how they are applied. They can do this by affecting the actions of parties to the criminal process, e.g., legislators, judges, police, and jail or prison staff. The aim to minimize harm also motivates development of less harmful enforcement methods that are comparably effective at achieving enforcement aims. This can amplify the differences over time. The aim to minimize harm also motivates efforts to reduce the need for enforcement, e.g., through social programs. To put it generally: rejecting the aim to harm and aiming to minimize harm motivates restraint and guides enforcement's development in ways calculated to reduce its harmfulness.

My response to the objection appeals to what T.M. Scanlon calls the "predictive significance" of intentions or aims (2008: 30–32). I think my response is decisive. Even if the Doctrine of Double Effect is false, there's an important difference between punishment and non-punitive enforcement of the sort I've described.

My response also highlights another difference between my argument and certain abolitionist arguments. Abolitionists are often unclear about how they view incidental harm. Geoffrey Sayre-McCord appeals to the Doctrine of Double Effect in defense of replacing punishment with reparations. But he recognizes that even if the doctrine is false, one could appeal to a predictive significance claim (2001: 507). He doesn't give an argument for such a claim, though. And he expresses ambivalence about what view reparative enforcement should take towards incidental harm (507, n14).

Boonin's treatment is similar. He appeals to the Doctrine of Double Effect repeatedly. Only once does he say that the aim to harm may affect enforcement's harmfulness (2008: 234). He considers a case where we have to decide whether to confine an offender in prison or at home. He says that punishment favors prison and that his proposed alternative—restitution—favors home detention. But this is dubious. Whether punishment favors imprisonment depends on how much and what kinds of harm it aims to inflict. And a non-punitive system that is indifferent to the incidental harms it inflicts could easily inflict more harm than an otherwise similar system of punishment (a point Sayre-McCord also overlooks). Philosophers typically defend punishments that aim to inflict only a certain amount of harm and no more. Indifference to harm imposes no such limit. As far as I can tell, Boonin neither endorses nor rejects such indifference.[18]

My response to the objection suggests obvious ways to modify IA, N, and D. One might worry that rejecting the aim to harm and aiming to minimize harm would compromise

[16] This claim is cautious. Stronger claims seem plausible, e.g., that the non-punitive system probably would harm offenders less. I won't press these claims. The weaker claim is all I need.

[17] This difference isn't the only potentially significant one. I think that rejecting the aim to harm and aiming to minimize harm expresses respect for offenders and a commitment to treating them humanely, for example (thanks to Thaddeus Metz for this suggestion). I can't defend these claims here. Elsewhere, I've argued against retributivist views. These arguments are relevant to this issue. See Hanna (forthcoming a, forthcoming b). Those arguments lend some support to these claims.

[18] He does say that if the point of incarceration is restitution "there will be no justification for making [the offender's] life any less pleasant than is required by his incarceration" (234). But it's not clear whether he's talking about incidental harm here or harm we're aiming at.

deterrence too much. I'll consider some modifications of the claims along these lines. I'll also examine contemporary deterrence research and argue that it doesn't support such claims.[19]

Some Modifications

As a first pass, consider this set of claims.

IA_1 Non-punitive enforcement cannot deter enough
N_1 Legal punishment is necessary to deter enough
D_1 The fact that legal punishment deters enough is one reason why it is justified

The problem here is that it's not clear what the proposed deterrent standard is (cf. Nozick 1974: 61). What counts as enough? Without more information, there's no obvious reason to believe IA_1 and N_1. The modification suggests an important strategy, though: identify a morally significant quantity of deterrence that punishment is necessary to secure.

Alternatively, one might grant that non-punitive enforcement can deter enough, but argue that punishment is necessary to deter in the right way. Retributivists might say that only punishment deters in a just way or in a way that gives offenders what they deserve. Expressivists might say that only punishment deters in a way that sends the right messages. I won't consider such views. They abandon the claim that punishment's deterrent capability helps to justify it over non-punitive enforcement. I'm concerned to challenge that sort of claim. I'll focus on modifications that appeal only to deterrence.[20]

Next, consider this modification, suggested by some of the philosophers discussed earlier.

IA_2 Non-punitive enforcement cannot deter enough to keep society stable, decent, or tolerable
N_2 Legal punishment is necessary to deter enough to keep society stable, decent, or tolerable
D_2 The fact that legal punishment deters enough to keep society stable, decent, or tolerable is one reason why it is justified

This modification says that punishment is justified partly because it's needed to achieve some threshold of stability, decency, or tolerability. This modification is an improvement, since the deterrent standard is clearer. But the claims aren't obviously true. There's no obvious reason to believe that non-punitive enforcement can't deter enough to meet the threshold, assuming there is one. These claims must be supplemented with details about the threshold and evidence that non-punitive enforcement can't meet it.

I'll discuss whether such evidence is available soon. For now, consider another modification that might avoid the need for it.

IA_3 Non-punitive enforcement cannot deter as much as legal punishment

[19] This also distinguishes my argument from Boonin's. He doesn't examine the deterrence literature or claims like those below—the one exception being the third set of claims (267). He may seem to discuss the first and second set, but this is because his wording is misleading. He argues that restitution can deter "sufficiently" (264–66). By this he means that it can make the expected costs of offending outweigh its expected benefits, and so can make it such that offenders rationally *should* be deterred (264). I agree, but this doesn't refute any of the claims below. This is because nothing follows about how many offenses *will be* deterred.

[20] For an argument against expressivism see Hanna (2008). For arguments against retributivism see Hanna (2009a, forthcoming a, forthcoming b).

N_3 Legal punishment is necessary to deter more than non-punitive enforcement
D_3 The fact that legal punishment deters more than non-punitive enforcement is one reason why it is justified

Advocates of this modification might claim that punishment's greater deterrent power can make society *more* stable, decent, or tolerable than non-punitive enforcement. This modification makes two assumptions: that non-punitive enforcement can't deter as much as punishment and that punishment's additional deterrent power helps to justify it. Even if we grant the first assumption, there's no obvious reason to believe the second. Even if punishment can deter more than non-punitive enforcement, and thereby make society more stable, decent, or tolerable, that doesn't obviously help to justify it. Suppose torture would make society more stable, decent, or tolerable than standard punishments. That's not a good reason to torture (cf. Boonin 2008: 267).

Deterrence justifications need a better set of claims. My remarks suggest some criteria that justifying claims must meet.

(i) They must state a clear deterrent standard.
(ii) There must be evidence that non-punitive enforcement can't satisfy the standard and that punishment can.
(iii) Satisfaction of the standard should be important enough to help justify punishment over non-punitive enforcement.

If contemporary deterrence justifications are any indication, these criteria are surprisingly difficult to satisfy—more difficult than advocates of these justifications think. To give a prominent example, Erin Kelly says that "the threat of punishment for criminal wrongdoing seems easy to justify to those who value the rights a system of punishment aims to protect" (2009: 449).[21] This significantly overstates the deterrent case. Next, I'll consider additional modifications of the claims. They face similar difficulties.

Risk and Deterrence Research

Another modification accommodates Farrell's concerns about risk.

IA_4 Non-punitive enforcement cannot deter enough to sufficiently reduce the risk of wrongful harm posed by offending
N_4 Legal punishment is necessary to deter enough to sufficiently reduce the risk of wrongful harm posed by offending
D_4 The fact that legal punishment can deter enough to sufficiently reduce this risk is one reason why it is justified

Even granting that these principles satisfy criterion (iii) (because of the word *sufficiently*), they don't obviously satisfy (i) or (ii).

Clarifying the standard to satisfy (i) isn't easy. Simplifying, suppose it could be put in terms of a general value or range of values R for all offenses and associated harms (more realistically, one might propose different sets of principles about different offenses or offense types).

[21] This claim is about the *threat* of punishment, but Kelly thinks that the threat is justified only if punishment is (2009: 450).

IA_5 Non-punitive enforcement cannot deter enough to reduce the risk of harm posed by offending by R

N_5 Legal punishment is necessary to deter enough to reduce the risk of harm posed by offending by R

D_5 The fact that legal punishment can deter enough to reduce the risk of harm posed by offending by R is one reason why it is justified

Attempts to defend such principles face two hurdles. First, one must show that punishment's alleged ability to secure R would help to justify it if non-punitive enforcement can't secure R. That is, we need a reason to think that criterion (iii) is satisfied—that for the proposed standard, the move from IA_5 and N_5 to D_5 is plausible. The need to defend such a move is typically overlooked because philosophers assume that non-punitive enforcement can't meet even the most minimal standards. As we've seen, there's no reason to accept this assumption.

Second, we need a reason to think that criterion (ii) is satisfied. That is, we need evidence for IA_5 and N_5. The need for such evidence is sometimes acknowledged, but it's usually overlooked in favor of quick arguments or assumptions like those we've seen. Again, this is because the misrepresentation of our options makes it seem obvious that non-punitive enforcement can't meet any plausible standard. Given our options, there's no obvious reason to accept this. We need empirical evidence that non-punitive enforcement can't satisfy whatever the right standard is.

There's no good empirical evidence for such a claim. A quick survey of contemporary deterrence research illustrates this. There's agreement in the literature that criminal justice systems deter significantly.[22] But that claim is consistent with the claim that non-punitive criminal justice systems can deter significantly. What deterrence justifications need is evidence for a claim like this:

> The deterrent effects of contemporary enforcement systems are due largely to their punitive character.

Deterrence research doesn't support this claim.

First, many deterrence researchers endorse claims that cast doubt on the above claim. They're quick to grant that things like policing strategies, perceived risk of apprehension, and informal effects triggered by formal sanctions all deter substantially (cf. Nozick 1979: 59–60). Non-punitive enforcement can deter by such means—and deliberately so. It can modify policing strategies, manipulate the perceived risk of apprehension, and publicize information about enforcement's incidental harms in order to deter more. And it can present such information in ways calculated to deter more. These are just a handful of ways to try to deter while using non-punitive enforcement.

Second, contemporary deterrence research isn't geared towards assessing the above claim. This is partly because deterrence researchers typically take the same mistaken view of our options that philosophers do. This precludes evaluation of the kind of enforcement I've discussed. For example, Gary Becker, in a classic and widely cited paper, just assumes that many standard enforcement techniques are punishments (1968: 179–180). Similarly, Daniel Nagin's (1998) research survey contains no discussion of the nature of punishment or of the extent to which deterrent effects are due to the punitive nature of contemporary enforcement. Nor does it suggest that these issues have been of concern to researchers.

[22] See, e.g., Doob and Webster (2003: 144), von Hirsch et al. (1999: 1), Nagin (1998: 3), and Robinson and Darley (2004: 173).

Ditto for another survey, whose authors explicitly define criminal deterrence as "the avoidance of criminal acts through fear of punishment" (von Hirsch et al. 1999: 5). The unfortunate implication here is that only punishment deters. The definition suggests that the authors mistakenly assume that coercive enforcement and punishment are identical. If so, they're assuming that the sort of enforcement I've described is either incoherent or just is punishment.[23]

Deterrence researchers usually focus on narrower issues like whether tougher sentences deter more.[24] To take one example, Steven Levitt's (1996) work on the relationship between prison population and crime rates purports to show that imprisonment can deter significantly. But he doesn't investigate the extent to which the alleged effect is due to the punitive nature of contemporary imprisonment. His (1995) work purporting to demonstrate similar effects from higher arrest rates also can't credit the effects to the punitive nature of contemporary enforcement.

I conclude that deterrence justifications rest on unsupported assumptions. Fundamental confusions obscure this and make these assumptions seem obviously true. Exposing the errors makes the lack of evidential support for the assumptions clear. Given the centrality of these assumptions, we must conclude that extant deterrence justifications fail.

Objections

I'll consider two objections that offer armchair defenses of the claim that non-punitive enforcement of the sort I've described can't deter enough.

Here's the first. Since non-punitive enforcement doesn't aim to harm, it can't fine-tune the harmfulness of sentences to deter enough. But it's very unlikely that we'll deter enough if we don't fine-tune the harm. Indeed, it's even more unlikely if we're trying to make enforcement less harmful. So it would be a big coincidence if non-punitive enforcement deterred enough. So punishment is probably necessary to deter enough. Call this the Unlikelihood Objection.[25]

Here's the second objection. There are offenses that non-punitive enforcement isn't equipped to deal with and won't be able to deter. The idea here is that the aim to harm is at least sometimes needed to motivate an enforcement strategy. Take a simple example: speeding. Speeders are usually fined. One might worry that rejecting the aim to harm leaves us with no reason to impose fines. And even if there are such reasons, one might worry that there's no way to motivate payment without threatening harm, and so no way to deter speeding. Similar things could be said about other offenses.[26] Call this the Unavailability Objection.

[23] Prominent commentators' claims about the state of deterrence research seem to support my claims about the research on non-punitive enforcement. Andrew Ashworth says that "there has never been thoroughgoing examination in [the U.K.] of … whether some [effective] form of non-criminal enforcement could be devised" (1995: 50–51). Douglas Husak says that this observation applies to the U.S. too (Husak 2004: 208, n3). We could just replace the term *non-criminal* here with *non-punitive* (given that neither Ashworth nor Husak distinguish these in the same way I do).

[24] For empirically informed skepticism about this issue see Doob and Webster (2003) and Robinson and Darley (2004).

[25] Thanks to an anonymous referee for this sort of objection.

[26] I take it one couldn't say such things of offenses like murder and rape, since we have strong reasons to incapacitate murderers and rapists and to make them furnish compensation. This isn't so obviously true of speeding. Many offenses might be like speeding in this regard.

Consider the Unavailability objection first. It's surprisingly unimaginative. Sticking with our example, there are lots of non-punitive ways to deal with speeding that are well motivated and that can deter. Many things we already do could be done non-punitively. Presumably, we'd want police to pull over speeders and censure them, if only to stop them and remind them that what they're doing poses unacceptable risks to others. This can deter. Instead of punitive fines, officers could issue compensation orders to speeders to defray enforcement costs. Arguably, speeders should be responsible for more of the costs. Those who can't pay could be required to provide non-monetary compensation, e.g., by cleaning up litter. Those who refuse to pay could have their wages garnished or some of their assets seized. More compensation could be required to defray the additional administrative costs in such cases.[27] Especially egregious offenses or repeat offenses could be met with license suspension/revocation or vehicle seizure, for purposes of incapacitation. Taking driver's education classes could be required—for purposes of reformation—before reinstating licenses or returning seized vehicles. Or it could be encouraged with positive incentives, e.g., by offering to reduce compensation orders or the length or invasiveness of incapacitation measures.[28]

This just covers *some* non-punitive strategies. It doesn't cover all the details or address every potential worry. But it hopefully shows that non-punitive strategies and plausible rationales for them are available. This is just one offense, though. One might worry that we'd have a much harder time with others. I can't discuss every such alleged offense here. All I can say is this: flatly asserting that there are offenses that non-punitive enforcement isn't equipped to deal with is too quick. Serious, sustained attempts to develop non-punitive strategies must be made first. The Unavailability Objection will be compelling only if such attempts fail (and even then, it will only apply to some offenses).[29] A critic's inability to easily come up with such proposals is hardly convincing. I've offered no such proposal here, partly for reasons of space, partly because it isn't necessary for my purposes, and partly because I don't think philosophers can competently formulate such proposals alone. Among others, criminal justice professionals must take part.

So consider the Unlikelihood Objection. It relies on questionable assumptions, only some of which I'll address. Among them are an assumption about how much deterrence is enough and an assumption about the deterrent effectiveness of non-punitive enforcement. I've already addressed such assumptions. But it's worth reiterating the problems with them. The assumptions are remarkably persistent. And the problems are easily forgotten or overlooked.[30]

[27] Since I've mentioned compensation several times, I must address an objection. One might object that there are offenses that can't be completely compensated for. This seems right, but it doesn't pose a problem for anything I've said. Keep in mind that all I'm doing here is challenging claims like D. Towards this end, all I'm saying about compensation is the following. It's a non-punitive technique that can deter. And we often have reason to make offenders furnish compensation even if we're not aiming to harm them. These claims are consistent with the claim that the objection appeals to.

[28] I hope the rationale here is obvious. But to spell it out: offenders who complete reformation programs may pose less of a risk, so less restrictive incapacitation may be warranted. And the long-term gains here might be worth the costs associated with reduced compensation orders.

[29] Note, however, that the absence of a plausible non-punitive enforcement rationale may speak against criminalizing the conduct and not against non-punitive enforcement.

[30] There's another questionable assumption: that the aim to harm lets us fine-tune deterrence. I won't assess this claim, but I think it overstates our control over deterrent effects. Punishment is a blunt instrument, not the surgical one this assumption takes it to be. Again, see Doob and Webster (2003) and Robinson and Darley (2004) for empirically informed skepticism here.

As for the first assumption, how much deterrence counts as enough is controversial. As I've said, any such standard must be stated and assessed. Once stated, we need arguments for the standard and evidence that non-punitive enforcement can't satisfy it. We can't just assume that, whatever the standard, non-punitive enforcement can't satisfy it.

As for the second assumption, it should be clear by now how easy it is to underestimate and misrepresent non-punitive enforcement's deterrent effects. Important errors contribute to this. Advocates of deterrence justifications underestimate the range of non-punitive options. And the Unavailability Objection suggests how easy it is to underestimate the prospects for developing non-punitive enforcement strategies. A big part of the problem here is that we're just not used to thinking about enforcement in these terms.[31] Given the complexity of these issues and the substantial room for error here, the armchair probability assessment that the Unlikelihood Objection relies on is doubtful.

But one might complain that I'm underestimating the unlikelihood objection—or part of it, at least. The sort of non-punitive enforcement I've described doesn't aim to harm and aims to minimize the incidental harms it inflicts in pursuit of its aims. One might ask: What if we can achieve those aims harmlessly? Or at least develop harmless methods? Either way, such enforcement would or could come to have little to no deterrent power. One might complain that I can't rule out these possibilities and that this poses a serious problem for my argument.[32]

Part of this is right. I can't rule out these possibilities. But that's not a problem for my argument. I've argued that extant deterrence justifications fail because they rely on unsupported empirical assumptions. The fact that I can't rule out these possibilities doesn't entail that this is false or even implausible. It just means that, so far as my argument goes, philosophers of punishment might be able to develop a successful deterrence justification. I haven't denied this. To do this, though, they'd need evidence that non-punitive enforcement of the sort I've described wouldn't deter enough. The fact that I can't rule out the aforementioned possibilities doesn't constitute such evidence.

Though I can't rule out these possibilities, there are reasons to doubt that they're realistic possibilities. It's not obvious how to effectively and harmlessly incapacitate the dangerous. Nor is it obvious how offenders could furnish compensation in ways that don't harm them at all.[33] Ditto for any number of other aims, e.g., reforming offenders and criticizing them.[34] Even if there are harmless, effective ways of achieving some enforcement aims, there may not be such ways of achieving others. None of this shows that there are no ways of harmlessly and effectively achieving such aims, of course. But it does show that there's no obvious, pressing problem for my argument here. That is, there's no problem here that vindicates extant deterrence justifications against my criticisms.

[31] Indeed, many philosophers have discussed our overreliance on punishment. See, e.g., Huemer (2012), Husak (2004, 2009).

[32] Thanks to an anonymous referee for pressing me here.

[33] Obviously, the harmfulness here will vary. Richer offenders, for example, will be harmed less by paying monetary compensation orders. But this doesn't show that a non-punitive enforcement system of the sort I've described would or could be harmless and/or unable to deter enough, let alone at all. Some of the methods it uses will deter some potential offenders less than others. But this is true for any enforcement system.

[34] Maybe some of these aims could be harmlessly achieved with highly advanced technology. But that possibility hardly brightens the present prospects for deterrence justifications. And I'm dubious that it brightens the future prospects for them either. A society with such technology may very well have technology (or the ability to easily develop technology) that makes deterring offenses unnecessary.

Conclusion

To repeat: Nothing I've said shows that deterrence justifications can't succeed. What I've argued is that extant deterrent justifications fail. There are important obstacles to formulating a successful deterrence justification. These obstacles should be acknowledged and taken more seriously. Contemporary deterrence justifications rest on unsupported assumptions about the nature and deterrent effectiveness of non-punitive enforcement.[35]

References

Ashworth, A. (1995). *Principles of Criminal Law* (2nd ed.), New York: Oxford University Press.

Barnett, R. (1977). "Restitution: A New Paradigm of Criminal Justice," *Ethics*, 87: 279–301.

Becker, G. (1968). "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, 76: 169–217.

Bentham, J. (1843). *Principles of Penal Law* in J. Bowring (ed.), *The Works of Jeremy Bentham*, vol. 1: 367–580, London: Simpkin, Marshall.

Berman, M. (2008). "Punishment and Justification," *Ethics*, 118: 258–290.

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

Bradley, B. (2009). *Well-Being and Death*, New York: Oxford University Press.

Braithwaite, J. (2002). *Restorative Justice and Responsive Regulation*, New York: Oxford University Press.

Doob, A.N. and Webster, C.W. (2003). "Sentence Severity and Crime: Accepting the Null Hypothesis," *Crime and Justice*, 30: 143–95.

Duff, R.A. (2001). *Punishment, Communication and Community*, New York: Oxford University Press.

Ellis, A. (2003). "A Deterrence Theory of Punishment," *The Philosophical Quarterly*, 53: 337–51.

Farrell, D. (1990). "The Justification of Deterrent Violence," *Ethics*, 100: 301–17.

Farrell, D. (1988). "Punishment Without the State," *Nous*, 22: 437–53.

Farrell, D. (1985). "The Justification of General Deterrence," *The Philosophical Review*, 94: 367–94.

Feinberg, J. (1963). "Justice and Personal Desert," in C. J. Friedrich and J. W. Chapman (eds.), *Nomos VI: Justice*, New York: Atherton Press, reprinted in J. Feinberg, 1970, *Doing and Deserving*, Princeton: Princeton University Press.

Golash, D. (2005). *The Case Against Punishment*, New York: New York University Press.

Hampton, J. (1984). "The Moral Education Theory of Punishment," *Philosophy and Public Affairs*, 13: 208–38.

Hart, H.L.A. (1959). "Prolegomenon to the Principles of Punishment," *Proceedings of. the Aristotelian Society*" 60: 1–26, reprinted in H.L.A. Hart, 1968, *Punishment and Responsibility*, New York: Oxford University Press.

von Hirsch, A., Bottoms, A., Burney, A., and Wikstrom, P-O. (1999). *Criminal Deterrence and Sentence Severity: An Analysis of Recent Research*, Portland, OR: Hart Publishing.

Hanna, N. (forthcoming a). "Retributivism revisited," *Philosophical Studies*.

Hanna, N. (forthcoming b). "Two Claims About Desert," *Pacific Philosophical Quarterly*.

Hanna, N. (2012). "It's Only Natural: Legal Punishment and the Natural Right to Punish," *Social Theory and Practice*, 38: 598–616.

Hanna, N. (2009a). "The Passions of Punishment," *Pacific Philosophical Quarterly*, 90: 232–250.

Hanna, N. (2009b). "Liberalism and the General Justifiability of Punishment," *Philosophical Studies*, 145: 325–49.

Hanna, N. (2008). "Say What? A Critique of Expressive Retributivism," *Law and Philosophy*, 27: 123–150.

Hoekema, D. (1980). "The Right to Punish and the Right to be Punished," in H.G. Blocker and E.H. Smith (eds.), *John Rawls' Theory of Social Justice: An Introduction*, Athens, OH: Ohio University Press, 239–69.

Huemer, M. (2012). *The Problem of Political Authority*, Palgrave Macmillan.

Husak, D. (forthcoming). "Retributivism *In Extremis*," *Law and Philosophy*.

Husak, D. (2009). *Overcriminalization*, New York: Oxford University Press.

Husak, D. (2004). "The Criminal Law as Last Resort," *Oxford Journal of Legal Studies*, 24: 207–235.

[35] Thanks to Ken Baynes, Ben Bradley, Edward McClennen, Thaddeus Metz, Michael Stocker, Bill Wringe, and several anonymous referees.

Husak, D. (1992). "Why Punish the Deserving?," *Nous*, 26: 447–464.
Kelly, E. (2009). "Criminal Justice Without Retribution," *Journal of Philosophy*, 106: 440–62.
Levitt, S. (1996). "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," *The Quarterly Journal of Economics*, 111: 319–351.
Levitt, S. (1995). "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" *National Bureau of Economic Research Working Paper 5268.*
Lucas, J.R. (1968). "Or Else," *Proceedings of the Aristotelian Society*, 69: 207–222.
Montague, P. (2002). "Recent Approaches to Justifying Punishment," *Philosophia*, 31: 1–34.
Montague, P. (1995). *Punishment as Societal-Defense*, Lanham, MD: Rowman & Littlefield.
Nagin, D. (1998). "Criminal Deterrence Research at the Outset of the Twenty-First Century," *Crime and Justice*, 43: 1–42.
Nino, C.S. (1983). "A Consensual Theory of Punishment," *Philosophy & Public Affairs*, 12: 289–306.
Nozick, R. (1979). *Anarchy, State and Utopia*, Basic Books.
Quinn, W. (1985). "The Right to Threaten and the Right to Punish," *Philosophy and Public Affairs*, 14: 327–73.
Rawls, J. (1971). *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Rawls, J. (1955). "Two Concepts of Rules," *The Philosophical Review*, 64: 3–32.
Robinson, P. and Darley, J. (2004). "Does Criminal Law Deter? A Behavioral Science Investigation," *Oxford Journal of Legal Studies*, 24: 173–205.
Sayre-McCord, G. (2001). "Criminal Justice and Legal Reparations as an Alternative to Punishment," *Philosophical Issues*, 11: 502–29.
Scanlon, T.M. (2008). *Moral Dimensions*, Cambridge, MA: Harvard University Press.
Tadros, V. (2011). *The Ends of Harm*, New York: Oxford University Press.
Zimmerman, M. (2011). *The Immorality of Punishment*, Ontario: Broadview Press.