

---

# THE PASSIONS OF PUNISHMENT

BY

NATHAN HANNA

---

**Abstract:** I criticize an increasingly popular set of arguments for the justifiability of punishment. Some philosophers try to justify punishment by appealing to what Peter Strawson calls the *reactive attitudes* – emotions like resentment, indignation, remorse and guilt. These arguments fail. The view that these emotions commit us to punishment rests on unsophisticated views of punishment and of these emotions and their associated behaviors. I offer more sophisticated accounts of punishment, of these emotions and of their associated behaviors that are consistent with Abolitionism, the view that punishment is unjustified.

## *Introduction*

Some philosophers argue that our emotional reactions to wrongdoing commit us to punishment. Things like resentment, indignation, remorse and guilt – what Peter Strawson (1974) calls the *reactive attitudes* – play a significant role in our lives. They also seem intimately bound up with punishment, often motivating us to punish or to submit to punishment. Because of this, some philosophers think these emotions hold the key to justifying punishment. The justifications take different forms, but they share common claims. Proponents of punishment claim that these emotions are significant and fundamental elements of our lives and that they call for or at least reveal the importance and value of punishment. Because of this, these theorists argue, failing to punish is unacceptable, perhaps because it shows inadequate respect for persons or because it constitutes inadequate recognition of the gravity of wrongdoing.

These arguments fail. I will argue that these emotions do not commit us to punishment and that they are consistent with Abolitionism, the view that punishment is unjustified.<sup>1</sup> The view that these emotions commit us to punishment rests on an inadequate view of punishment and of these emotions and their associated behaviors. Specifically, philosophers have mis-

understood both the role that harm and suffering play in punishment and the role they must play in our everyday interactions with one another. Philosophers have consequently overestimated the importance that these emotions assign to punishment.

I begin by outlining some representative arguments for punishment that focus on these emotions. I then discuss punishment and clarify the role that harm and suffering play in punishment. With a clearer conception of punishment in hand, I then criticize these arguments. In the process, I offer a partial account of the emotional and moral significance of harm and suffering in the context of wrongdoing. This account demonstrates the consistency of Abolitionism with the reactive attitudes and their associated behaviors. I argue that harm and suffering may be of significant importance in the context of wrongdoing, but not for reasons that favor punishment.

### *1. From passions to punishment*

The arguments that concern me take their inspiration from Strawson's classic discussion of punishment and the reactive attitudes (Strawson, 1974). Strawson is concerned with issues of free will and determinism and their bearing on punishment's justifiability. He argues that punishment may be justified even if determinism is true. If we attend to 'that complicated web of attitudes and feelings which form an essential part of the moral life,' he claims, a justification of punishment can be formulated that takes no stand on the truth of determinism (Strawson, 1974, p. 23).

Determinism does not concern me, so I will not mention it again. Punishment and its relations to the reactive attitudes are what concern me. To put Strawson's argument simply, he observes that we have an ineliminable commitment to interpersonal relationships. Being involved in these relationships, we are prone to various attitudes, feelings and associated behaviors. These attitudes and feelings include resentment, indignation, remorse and guilt, which are part of a set of natural reactions to other people's attitudes, attitudes expressed by their behavior. When people behave in ways that express indifference to or ill will towards others, our attitudes and feelings towards them change. We do not have the goodwill we might otherwise have had towards them. We withdraw our goodwill or are no longer prone to extend them goodwill. We resent them and feel indignant. We blame and criticize them. We expect them to feel guilt and remorse. And sometimes we punish them.

Because of these feelings we – victims, third parties and offenders themselves – are often willing 'to acquiesce in that infliction of suffering on the offender which is an essential part of punishment' (Strawson, 1974, p. 22).

Withdrawal of goodwill and the willingness to have offenders suffer is part of what it is to have these feelings. The tendency to have and to express these feelings is part of what it is to be in interpersonal relationships with others. A wholesale suspension of or refusal to express these attitudes is, Strawson thinks, simply impossible. It would lead to treating others simply as 'object[s] of social policy' rather than as participants in interpersonal relationships (Strawson, 1974, p. 9).

Strawson thinks these observations say something important about what an acceptable justification of punishment will look like. He acknowledges that punishment may be partially justified by its good consequences. But he suggests that a justification must acknowledge the importance of the reactive attitudes and the role punishment plays in expressing them. Punishment cannot simply be a way of manipulating people. There has to be more to it than that. That something more is the expression of the reactive attitudes.

Christopher Bennett (2002) finds Strawson's account compelling and offers a retributivist argument for punishment that appeals to these attitudes.<sup>2</sup> By examining them and the role they play in informal everyday situations, he aims to demonstrate our commitment to retribution and punishment.

[O]ur participation in the reactive attitudes betrays commitment to retribution, to the thought that it is non-contingently a good thing that those who have done wrong should undergo certain forms of suffering. (Bennett, 2002, p. 147)

He tells a story about a man named Bryson. Bryson comes to work one day to find his coworkers giving him scornful looks and refusing to talk to him. After some consideration, Bryson realizes that it is because he has been neglecting his responsibilities and because he has been cheating on his partner Kate, who is friends with some of his colleagues. Bryson feels guilty, realizes he has treated everyone badly and decides that he deserves their scorn and more. He resolves to take his responsibilities more seriously and hopes that Kate will let him apologize.

Bennett tells this story to illustrate the importance of certain forms of suffering that we want wrongdoers to experience. He thinks these forms of suffering have considerable meaning and moral significance (Bennett, 2002, p. 148). There are three forms of suffering whose importance the story is meant to highlight.

First, there is the suffering involved in lost solidarity, the suffering Bryson experiences when his coworkers withdraw from him and treat him in ways that express their disapproval and blame. Being excluded is painful for social creatures like us. When people treat others poorly, we naturally withdraw from them. We alienate them because they have alienated themselves from important values.

Second, there is the suffering involved in repentance. It hurts to acknowledge the fact that one has wronged others. When Bryson's coworkers blame him, they are, among other things, trying to compel him to face up to what he did. He then comes to share their disapproval. He also feels guilt and remorse. Feeling this way about oneself and one's actions is painful.

Third, there is the suffering involved in reconciliation. We expect wrongdoers to apologize and to make amends. Bryson realizes that he has to stop shirking his responsibilities and he recognizes the need to apologize. Moreover, he should come to realize that he owes everyone more than mere apologies. He owes some debts. He could compensate his coworkers somehow, perhaps by offering to take on extra responsibilities. But we may think something more is in order, perhaps for his coworkers and certainly for Kate, in whose case compensation does not even seem appropriate. Bryson should, Bennett observes, perform some sort of penance to show how sorry he is, something done to benefit those he offended.

Whatever forms these gestures take, Bennett points out, they will all involve suffering. Apologizing is embarrassing. Compensating others and performing penances involve self-sacrifice. These forms of suffering are important because we often take a wrongdoer's willingness to undergo them to be an indication of sincerity. Wrongdoers who claim to realize the error of their ways without doing anything more to show their regret belie their claims. We rightly suspect that they are not really sorry or that they are insufficiently sorry for their wrongs.

Bennett thinks that the importance of these forms of suffering underscores the superficiality and unfairness of one popular view of retributivism. On his view, Retributivism is not about inflicting indiscriminate physical suffering and it need not be motivated by vengefulness or sadism. Our ordinary reactive attitudes – Bennett thinks it would be better to call them *retributive attitudes* – commit us to Retributivism. Examining our reactions in informal everyday situations like Bryson's helps us understand what we are doing when punishing. It may also help us discover how we should punish and how we should set up institutions of punishment (Bennett, 2002, p. 147).

Bennett (2002) does not discuss formal state punishment, but Antony Duff (2001) offers a helpful discussion of this issue that seems similarly inspired (cf. Bennett, 2008). He argues that:

[p]unishment should be understood, justified, and administered as a mode of moral communication with offenders that seeks to persuade them to repent their crimes, to reform themselves, and to reconcile themselves with those they have wronged. (Duff, 2001, pp. 115–16)

He rejects consequentialist justifications on the grounds that they show offenders and potential offenders insufficient respect. On Duff's view,

justified punishment is a way of communicating with people as responsible moral agents rather than a way of manipulating them. He takes the importance of communication to be obvious. We are naturally disposed to blame and criticize offenders. We want them to acknowledge their wrongs, to resolve to refrain from wrongdoing and to make amends. Censuring them gives them reason to do these things. From this foundation, Duff hopes to formulate a justification of punishment. He takes it to be obvious that offenders deserve censure. He argues that punishment is a necessary means of communicating the censure that serious offenses call for (Duff, 2001, p. 29).

Duff thinks we should try to get offenders to repent, to reform and to reconcile themselves with victims and the community. To do this, he argues, we must punish. Like Bennett, Duff thinks that focusing on reactive attitudes like resentment and guilt and on associated behaviors like criticism and making amends highlights the importance of certain kinds of suffering. They hold that there are important kinds of suffering we should try to impose on offenders. According to them, these kinds of suffering provide insight into punishment's justification.

When we censure offenders, we want them to repent their wrongs. But repentance, Duff observes, is 'necessarily painful, since it must pain me to recognize and admit . . . the wrong I have done' (Duff, 2001, p. 107). Since offenders deserve censure, they also deserve the suffering of repentance. Duff argues that punishment is necessary to generate repentance in cases of serious wrongdoing. This is because it is a forceful way of censuring. Verbally expressing censure, for example, is insufficient in such cases (cf. Duff, 2001, p. 82). Punishment also provides a structure within which offenders can focus on their wrongs and hopefully come to repent them (Duff, 2001, p. 108). Because genuine repentance is not something that can be achieved in a moment, such a structure is necessary to get offenders to focus on and contemplate their offenses. Genuine repentance leads to reformation. Reformed offenders resolve not to engage in further wrongdoing.

Repentant offenders, Duff observes, seek reconciliation with those they wronged. Reconciliation requires apology. Cases of serious wrongdoing require more than mere verbal apology, however. They require sufficiently forceful apologies that express sufficient repentance. Punishments constitute formal apologies of this sort and are therefore capable of at least partially reconciling wrongdoers and victims (Duff, 2001, p. 109). Reparation or other burdensome work performed for others' benefit sometimes suffices. But, Duff thinks, 'any kind of penitential burden' capable of expressing sufficient repentance will do (Duff, 2001, p. 109). The point is that offenders have to suffer some sort of burden to pay for their wrongs. Because of their wrongs they must somehow be weighed down. Offenders who claim to repent their wrongs but who suffer no apparent adverse

effects belie their claims. Offenders who think mere verbal apologies suffice for serious wrongs demonstrate either indifference or ignorance. People should feel badly when they do bad things. Offenders have to suffer. Something is wrong when they do not.

The theorists canvassed here think that informal observations of this sort hold in store crucial insights about punishment's justifiability. Certain forms of suffering are important and offenders should be made to experience them. They think a justification of punishment should start with such insights. On this view, punishment is an intuitively attractive and necessary way of imposing certain important kinds of suffering. Our informal everyday reactions to wrongdoing are supposed to illustrate this.

Strawson does not try to explicitly justify punishment, but he thinks his observations help to defend punishment from certain criticisms. Bennett and Duff, however, are more ambitious. They offer arguments for punishment that owe much to Strawson. Their arguments rest on an inadequate view of punishment and of the emotions and associated behaviors discussed, however. In what follows, I will clarify the nature of each and argue that these emotions and their associated behaviors do not commit us to punishment.

## *2. Passions, punishment and suffering*

To evaluate these arguments, we must get clear on the nature of punishment and of the emotions and associated behaviors at issue. First, consider punishment. I cannot offer a comprehensive account here. Nor need I. There is one crucial characteristic of it I will focus on. When we punish people, we are, among other things, out to harm them or make them suffer. Harm and suffering are not incidental side effects of punishment. Punishments are imposed and designed, at least in part, *in order to* harm offenders or make them suffer.<sup>3</sup> Many philosophers agree, including Duff.

Punishment aims to inflict something painful or burdensome on an offender for his offense. [. . .] Nor are this pain and this burden mere unintended side effects of a procedure which is not designed to be painful or burdensome. (Duff, 1992, p. 49)

It is an essential and intended element of punishment . . . that the victim be made to suffer, and of liability that he be made to pay; these are not mere regrettable derivatives of the undertakings, but rather their [goals]. (Feinberg, 1963 [1970], p. 67)

I will not argue in detail for the claim that punishment aims to impose harm or suffering upon offenders. I do so elsewhere (Hanna, forthcoming; Hanna, 2008; Hanna, ms.). There are two points I want to make, however.

First, this partial conception is intuitive. It also helps make crucial distinctions between and explains important differences between punishment and practices like involuntary psychiatric commitment and pretrial detention (cf. Ten, 1987, pp. 14–15). These practices are often conducted in the context of enforcement and they also involve harm and suffering. The role harm and suffering play in these practices is importantly different from the role they play in punishment, however. One difference is that the harm and suffering involved in these practices is (at least ideally) not the product of design. On the contrary, efforts are or at least can be made to significantly minimize them. The harm and suffering involved in punishment, however, is to a significant extent the product of design.

Second, this partial conception is compatible with other conceptions that one might first consider rival conceptions. To take an example that is particularly important here, Joel Feinberg argues that the expression of criticism is essential to punishment (Feinberg, 1965). Punitive imprisonment, for example, has a critical character that involuntary psychiatric commitment lacks. One may want to reject my conception in favor of such a conception. The philosophers with whom I am concerned apparently accept Feinberg's conception (without, I should emphasize, explicitly rejecting mine). But my conception is compatible with his and is arguably needed to supplement it. We can and often do criticize offenders by treating them in ways designed to harm them or make them suffer. Arguably, this is a significant part of the basis for the differing critical characters of punishment and things like involuntary psychiatric commitment.<sup>4</sup>

Keep in mind that I am not offering a comprehensive account of punishment. I only claim that the aim to impose harm or suffering is essential to it. There may be other essential aims and there may be other aims often involved in particular punishments. Whether there are other aims and what they might be is not my concern here. We must remember, however, that many aims are compatible with the aim to impose harm or suffering. Simply pointing to other possible aims is not sufficient grounds for rejecting my conception given its intuitive appeal and its conceptual and explanatory utility.

Those who disagree with my claim that the aim to impose harm or suffering is essential to punishment can take what I have to say about punishment to be about a type of punishment. Taking my argument in this way, however, is not grounds for thinking it less significant. Even if we confine ourselves to talk of types, this type of punishment is the paradigmatic kind that philosophers and legal theorists are most concerned to justify. Understood in this way, my argument can be taken to highlight important limitations to Bennett's and Duff's arguments.

The relevance of this aim is obvious. If an action is performed without the aim to impose suffering or harm, then it is not punishment. This will be so even if it does cause suffering or harm. The philosophers with whom I

am concerned try to justify punishment by focusing on certain emotions and associated behaviors, emotions and behaviors that often cause suffering and harm. Feelings like resentment and indignation motivate blame, censure, withdrawal and so on. Being treated in these ways is painful. But being treated in these ways also causes painful feelings like repentance, remorse and guilt. These feelings motivate things like apology, reparation and penance, all of which involve suffering and harm.

In this complicated interplay of emotions and behaviors, theorists like Bennett and Duff find reason to conclude that certain forms of suffering are necessary and valuable. Given the role that suffering and harm play in punishment, there is a seemingly compelling similarity between it and these everyday interactions, interactions it seems we cannot do without. These philosophers exploit this similarity on punishment's behalf. But their conclusions can be forcefully resisted. If the aim to impose suffering or harm plays no essential role in these interactions, the prospects for justifying punishment on the basis of them is not as promising as these philosophers think. And if we pay close attention to the suffering and harm involved in these interactions, we may find that, even if they are important, they may be important for reasons that do not favor punishment.

Resisting these conclusions, however, requires more clarification. We need to examine some of the emotions and behaviors at issue and clarify their relation to suffering and harm. Again, a comprehensive account is beyond the scope of this paper, but some preliminary observations and distinctions will suffice.

Consider guilt. Much is made of the fact that we want wrongdoers to feel guilty. They often come to feel guilty when we blame them. But what is guilt? As a crude first pass, I suggest that there are two salient elements to guilt – or at least to the paradigmatic guilty state of mind (I set aside phenomena like survivor guilt). A guilty person has certain beliefs about herself and her acts: that those acts were wrong, that she is blameworthy for performing them and so on. There are also certain emotions involved, emotions directed at herself and her acts. Someone who feels guilty feels badly about things she has done and feels badly about herself for having done them.

These feelings and their behavioral manifestations are crucial indicators of sincerity and concern. When an offender claims to believe that her act was wrong and claims to be sorry, her claims will be more credible if she seems pained by what she did. These feelings indicate sincerity. But they also indicate something else: a certain kind of concern. An offender who is pained by her wrongful act demonstrates the right sort of moral concern. Being pained in this way shows that she cares about doing what is right and about not doing what is wrong. Unlike a sociopath or a devil (we might say), she is not indifferent to or pleased by wrongness. Rather, it is something that troubles her.



Consider these observations in light of the fact that the aim to impose suffering or harm is essential to punishment. Must we blame offenders to harm them or make them suffer? Given that there are several elements involved in guilt, we may be able to blame without doing this, even if blaming does cause guilt. Furthermore, the fact that the suffering of guilt is an indicator of sincerity and concern may account for why guilt is desirable and important – perhaps even good on the whole – and why we might want offenders to feel it. This is consistent with the claim that suffering itself is a bad thing that we should not aim to impose, however. It may be good for offenders to feel guilty, despite the badness of the suffering involved, because of (among other things) the sincerity and concern from which the guilt and suffering stem. This suggests a more complex picture of guilt that, I will argue, does not obviously support punishment.

Similar things can be said about the suffering involved in apology and reparation. This may not speak in favor of aiming to impose these or any other kinds of suffering or harm on offenders, and hence may not speak in punishment's favor. In the next section I will elaborate on these points and extend them to the remaining reactive attitudes and their associated behaviors. An adequate account of these attitudes and behaviors demonstrates that they do not commit us to punishment.

### 3. *Alternative accounts*

I have argued for some important claims about punishment and some of the reactive attitudes and associated behaviors. The aim to impose suffering or harm is essential to punishment. Guilt is comprised of certain beliefs as well as painful feelings. These feelings are indicators of one's sincerity and concern. I suggested that this account of guilt and the suffering that it involves might show that guilt, and the role it plays in our everyday interactions, does not speak in punishment's favor. In this section I will try to demonstrate this. I will offer similar accounts of blame, apology and reparations that are equally unhelpful to punishment's advocates.

I will begin with blame and its associated behaviors, move on to guilt and conclude with apology and reparations. When we blame someone, we disapprove of her or her actions. When we disapprove of something, we tend to express our disapproval. We can do this in different ways. We can do it verbally, just by saying what we think. But we can also do other things. We can express disapproval by means of scornful looks, withdrawal and even with punishment. Being subject to disapproval hurts. But must we aim to harm someone or make him suffer when we express disapproval? No. Certainly we can express disapproval in such ways, but we need not. We can tell a third party what we think, for example. That

seems like a genuine expression of disapproval. But we can also express our disapproval directly to the person of whom we disapprove without such an aim. We can, for example, express disapproval in careful, measured ways and even offer comfort to those we criticize. We often do this with children and close friends, for example (cf. Hanna, 2008, pp. 144–45).

This is not to say that expressions of disapproval will be guided by an aim to impose suffering or harm unless they are expressed in these ways. The examples are only meant to press the point that such an aim is not essential to expressions of disapproval. What seems essential is the expression of certain thoughts and attitudes. We might want to say that some act is wrong or that some wrongdoer is blameworthy. We might want people to know how we feel about some wrong or wrongdoer. But we do not have to aim to impose suffering or harm on anyone to do these things. Expressions of disapproval often cause harm and suffering. There are nevertheless good reasons to express disapproval. This alone does not show that we are committed to aiming to impose harm or suffering. If it did, we could similarly conclude that doctors are committed to such an aim because they often have reasons to use treatments with painful side effects. Unless it can be shown that the harm and suffering that often result from expressing disapproval are reasons to express it, it seems that the need to express disapproval does not speak in punishment's favor.

But one might object that I have overlooked something. We are concerned here with the expression of a special kind of disapproval, namely moral disapproval. And one might reasonably point out that we are committed to expressing not just some moral disapproval but adequate moral disapproval. Perhaps the aim to impose harm or suffering is sometimes required to do this. Sometimes it is not enough just to say that some act is wrong. It will not do in cases of murder, for example. And it will not do for Bryson's betrayal of Kate. Many think murderers deserve to be harmed, say by being imprisoned. Perhaps Bryson deserves to be harmed too, say by being shunned or abandoned by Kate and his friends. Few reactions seem more natural to us, after all.

In response to this objection one could ask what makes moral disapproval so special. Why think that adequate moral disapproval ever requires the aim to impose harm or suffering? Merely subjecting a murderer to verbal criticism does seem inadequate, but it may be so for reasons that have nothing to do with the need to express adequate moral disapproval. Murderers need to be incapacitated, for example. That may be one reason why simply criticizing a murderer is inadequate, and it is a reason that has nothing obvious to do with expressing moral disapproval. Why think that there are degrees of disapproval that are beyond our resources to express verbally?

This response is overly simplistic and subject to a plausible objection. What counts as sufficient depends on many things. Sometimes words are

not enough. Even if we forcefully condemn a murder as wrong, we risk not being taken seriously if that is all we do, for we would probably do more if we really thought it seriously wrong. Similarly, if Bryson's coworkers continued to be as friendly with him as before and only expressed disapproval verbally, they could reasonably be suspected of not thinking his actions all that bad. Our words must be backed up by actions of the right sort. One might argue that we have to aim to impose suffering or harm on some offenders or risk sending the wrong message.

We should agree with this objection to a point. Expressing adequate disapproval sometimes requires more than words because our words can be belied by other things we do. It does not follow from this that we must aim to impose suffering or harm, however. We can express disapproval in more forceful ways without so aiming.

Consider confinement. One might think that only confinement (or something comparably serious) can express adequate disapproval of murder. Presumably, this is because the alternatives are expressively inadequate. Putting murderers on probation or criticizing them verbally, for example, would suggest that murder is not terribly serious. Now, these claims may be true. But they do not show that the aim to impose suffering or harm is necessary to express adequate disapproval of murder. This is because we need not confine people in order to harm them or make them suffer. We can and do confine people without so aiming. Involuntary psychiatric commitment (at least in its ideal form) is one among several examples that illustrate this. We can confine people in order to hurt them, but we need not confine them for that reason. We need not confine to punish. This is one reason why there are (or at least can be) such drastic differences between punitive imprisonment and practices like involuntary psychiatric commitment.

Now, I do not mean to suggest that we do confine the mentally ill in order to express disapproval of them, their conditions or even their actions. Unlike wrongdoers, who can (setting skeptical worries aside) be faulted for their wrongs and their status as wrongdoers, we do not hold the mentally ill at fault for their actions or for being mentally ill.<sup>5</sup> The example is simply meant to show that we can confine people without aiming to harm them or make them suffer. Confinement need not be punitive. Non-punitive confinement can nevertheless be a serious response to wrongdoing – far more serious than the expressively inadequate alternatives considered above. Because of this, it is far less plausible to say that confining murderers non-punitively would, like probation or mere verbal criticism, belie our claims about the wrongness of murder. It seems, then, that this response is capable of expressing strong moral disapproval despite the lack of an aim to impose suffering or harm.<sup>6</sup> It is far more serious than probation or mere verbal criticism and use of it need not be limited to cases where we hold the prospective detainee faultless. More-

over, use of it is consistent with the view that rightfully imprisoned wrongdoers deserve suffering or harm. They may simply deserve the incidental suffering or harm such a practice will likely inflict. We need not take this fact, however, to be a reason to aim to impose suffering or harm on them. Advocates of punishment have no monopoly on desert claims.

The things I have said about confinement generalize. Similar things can be said about many of the techniques employed by contemporary criminal justice systems. Community service, compensation orders, probation and the various restrictions that can accompany probation need not be guided by an aim to impose suffering or harm and they can all express moral disapproval. Such techniques are often employed for a variety of reasons, among them the need to incapacitate dangerous offenders and the need to compensate victims. The aim to impose suffering or harm does not seem essential to any of these techniques and they can be employed without it – though in the absence of this aim, they could be employed quite differently.<sup>7</sup> These techniques, in short, are not necessarily punishments. Punishment has no monopoly on them. So it seems we have a wealth of techniques available for dealing with wrongdoing and expressing moral disapproval of it that need not be guided by an aim to impose suffering or harm.

Similar things can be said about everyday responses like withdrawal. Theorists like Bennett think responses like these speak in punishment's favor. But are these responses even all that much like punishment? No – or at least not essentially. We can withdraw without aiming to harm someone or make her suffer. There are good reasons for withdrawal that do not involve this aim. Often we withdraw simply to avoid getting hurt rather than to hurt. Kate could withdraw from Bryson, say by ending their relationship, because being around him stirs up painful memories and emotions. Or she could withdraw because she can no longer trust him. His coworkers could withdraw for similar reasons, though they would probably withdraw to a lesser degree. Being involved in personal relationships makes us vulnerable. We accept this vulnerability in part because of how rewarding personal relationships can be. But sometimes people act in ways that show them to be untrustworthy or that make the mere fact of being around them painful. Being in a personal relationship with someone untrustworthy is risky. Withdrawing to varying degrees can reduce the risk. Being in a personal relationship with someone it is painful simply to be around is not rewarding – quite the opposite. Withdrawal is a means of escape in such cases.

We can certainly withdraw in order to hurt people, and I do not deny that we sometimes do. The point is merely that we need not. The aim to impose suffering is not essential to withdrawal, as it is to punishment. So the importance of withdrawal does not obviously speak in punishment's favor. More must be said. Perhaps there remains a roundabout way of making the case, though. I have suggested some good reasons we might

have for withdrawal, but there may be others. Certainly the ones I gave are not the whole story. Perhaps we withdraw to make wrongdoers feel guilty and to get them to apologize and make amends. More generally, we tend to criticize wrongdoers so they will react in such ways. If what Bennett and Duff say about these things is right, perhaps some hope remains for their case.

So let us examine each in turn. I have already discussed guilt and why it may be good for wrongdoers to feel guilty despite the fact that guilt involves suffering, which we have as yet been given no reason to think good. To restate things briefly, there are at least two salient features of a paradigmatic guilty state of mind: various beliefs about the wrongness of one's acts, one's blameworthiness and so on and unpleasant feelings accompanying those beliefs. I suggested that these feelings might be evidence that someone has these beliefs and has an appropriate kind and degree of moral concern. I also suggested that the complex nature of guilt, the fact that it involves several different elements, complicates and may ultimately frustrate attempts to justify punishment on the basis of it. Guilt may be able to play an important social role without anyone aiming to generate the unpleasantness of the feelings involved in it. This is because there are other elements of it that we can and perhaps should aim at. The unpleasantness could be an incidental result of other important aims, one that we can do much to minimize.

What sorts of aims, though? An obvious candidate is an aim that seems central to blame and criticism: the aim to get offenders to view their acts as we do, i.e. as wrong.<sup>8</sup> There are a number of reasons to do this. Believing such things may make it less likely that offenders will repeat their behavior and can motivate them to make amends. Offenders may suffer if they adopt these beliefs, and their suffering may be a crucial indicator of our success in getting them to adopt these beliefs. The point, however, is that we need not try to get them to adopt the beliefs in order to harm them or make them suffer. There are other good reasons to get them to adopt these beliefs, reasons that can account for the importance of blame and guilt and that can justify acts of blame despite the suffering that may result. We can account for the importance of these beliefs in a way that does not speak in favor of an aim to impose suffering or harm. Because the account does not speak in favor of this aim, it does not speak in punishment's favor.

A note of caution before proceeding: I have said that the suffering involved in guilt can be an incidental result of actions aimed at doing things other than imposing suffering or harm. All I mean by this is that the suffering can be incidental *relative to certain aims*, e.g. an aim to generate certain beliefs. I do not mean to suggest that such suffering is incidental *simpliciter* (whatever that might mean) or trivial or unimportant. As I have said, it may be an important indicator of sincerity and concern. I do not even mean to suggest that we are, at least under normal conditions, crea-

tures who are capable of acquiring the beliefs in question without being pained by them. My only point is that the suffering is distinct from various other things, e.g. certain important beliefs, and that aiming to get offenders to adopt such beliefs need not involve an aim to harm them or make them suffer.<sup>9</sup>

One issue remains. As the philosophers I have discussed point out, we often expect wrongdoers to apologize and make amends. Doing these things is often painful. Apologizing can be embarrassing, for example, and furnishing reparations typically comes at a cost. In fact, we may think that reparations must sometimes come at a significant cost to be sufficient. Ways of making up for one's wrongdoing that require little to no effort or sacrifice seem unacceptable. Perhaps, one might argue, this is evidence that the suffering is valuable in itself. Perhaps like the very interactions in which it occurs, the suffering is a crucial and indispensable element in human relationships. Perhaps this is also the best way to understand punishment and the harm and suffering it aims to inflict.

Suffering may play an important role in human relationships and it may be an important part of making amends, but we need to know how and why it is important. I have given an account of the suffering involved in guilt that does not speak in punishment's favor. I offer a similar account of the suffering involved in making amends.

As with the suffering involved in guilt, we can say that the harm and suffering involved in making amends is an indicator of sincerity and moral concern. This observation may not be enough, however, for the harm and suffering involved in making amends is importantly different from that involved in guilt. The latter arguably has an involuntary character. It is partly an emotional reaction brought on by certain beliefs and attitudes. The harm and suffering involved in making amends, however, is voluntarily accepted insofar as the acts performed are voluntary. Perhaps we can in some sense be said to aim at the suffering and harm involved in making amends since we expect offenders to perform these actions at some cost to themselves. Perhaps even offenders can be said to aim at this suffering and harm because they accept it.

Nothing essential to making amends commits us to saying such things. Suppose repentant wrongdoers do accept the costs of making amends. Does this show that, in making amends, they must aim to harm themselves or make themselves suffer? Must they make amends in order to do these things to themselves? No. For they could make amends, not because of any suffering or harm it may cause them but *in spite of* any such suffering or harm. And the willingness to make amends despite the harm and suffering incurred can have powerful expressive value. It can serve to show that the offender is truly sorry for what she did, that she sincerely cares about righting her wrong and so on. A repentant wrongdoer could take the following view. She could see the suffering or harm as bad but nevertheless

worth the cost, given the benefits she can confer on her victims and the fact that she is obliged to compensate her victims.

What goes for wrongdoers can go for others. We can look on the suffering or harm involved in the same way. The suffering and harm may not be valuable in themselves, but they can be crucial indicators of sincerity and concern, and the willingness to undergo them in pursuit of the appropriate ends can be a powerful expression of apology. We expect wrongdoers to go to some effort to make amends because this gives us some assurance that they are sorry for their wrongs and that they want to make up for them. Empty acts easily performed are worth little, just as empty words easily mouthed. When a wrongdoer goes to some difficulty to make amends, we have good reason to believe that she really is repentant, really understands what she did, and is making amends for the right reasons. The willingness to incur harm and suffering upon oneself in service to others is often a mark of good character.

This is no reason to think the harm and suffering involved good, however. There are familiar contexts where similar things can be said. The harm and suffering experienced by saints and humanitarians are not themselves good. It is the saints and humanitarians who are good, in part because they are willing to make sacrifices in service to others. Repentant offenders typically are not saints or humanitarians (two important differences: the latter are not obliged to benefit others as they do, nor must they be at fault for some wrong). But similar things can be said about their self-sacrifice. The harm and suffering incurred may not themselves be good. But it is sometimes good to make sacrifices like this for others and such sacrifices can reflect well on those who make them.

This account of making amends puts it in a light that is unhelpful for advocates of punishment like Bennett and Duff. There are reasons to think that the harm and suffering involved are important. There are also reasons to think that certain acts (like making amends or apologizing) and certain mental states (like guilt) in which suffering and harm play a prominent role are good. But these are not reasons to think that suffering or harm are sometimes valuable in themselves. Nor are they reasons to think that suffering or harm are things we should aim to impose on offenders. If this account of making amends is plausible, then like my accounts of blame and guilt, it shows that the role that making amends plays in personal relationships does not commit us to punishment. We can make amends and expect others to do so without aiming to impose suffering or harm. Abolitionism can accommodate these emotions and behaviors – or at least their essential elements.<sup>10</sup>

I do not mean to suggest that the alternative accounts I have given of these interactions reflect how people actually behave. All I have argued is that there are no essential aspects of these interactions that commit us to punishment. Advocates of punishment will understandably feel uneasy

about many of the things I have said, however. Even if I am right about the ways we could act, think and feel, the fact is that we do not act, think and feel in these ways. We want to hurt wrongdoers. We often feel angry and vindictive, act on those feelings, and try to hurt them. One might think that this is evidence that we are committed to punishment and that punishment is justified. By way of conclusion, I want to briefly address lingering unease about the things I have said by sounding a note of skepticism about any attempt to justify punishment like those discussed.

Our allegiance to punishment runs deep. Practically everyone thinks that it is permissible to try to hurt certain offenders. In this respect, philosophers are not special. Philosophers disagree about why punishment is justified, but few question the claim that it is. Practically everyone thinks it is. Practically everyone wants to punish offenders.

These facts significantly influence our interactions with one another. Given this, to look to our actual interactions for a justification poses a significant risk of question begging. Punishment's advocates try to avoid begging the question by arguing for certain necessity claims: punishment is necessary to secure important goods, to communicate certain things, etc. Necessity claims like these are difficult to defend in light of the partial conception of punishment I have offered. This conception highlights one essential characteristic of punishment and, in doing so, highlights the possibility of substantive alternatives to punishment.

The theorists I have discussed formulate necessity claims with an eye on human relationships. They take certain aspects of the ways we think, feel and act – aspects they take to speak in punishment's favor – to be essential, to be indicative of how we must think, feel and act. I have argued that there are no essential elements of these interactions that speak in punishment's favor. Perhaps harm and suffering are in some sense essential and important aspects of the interactions at issue here, but that in itself is no reason to think punishment justified. Harm and suffering are ubiquitous. That is no reason to think them good, let alone a reason to think the aim to impose them good.

To make good the arguments considered, punishment's advocates must show that this aim is essential to human relationships, for this aim is essential to punishment. It is largely what makes punishment so difficult to justify. But this aim is arguably not essential to human relationships. It is certainly not essential to the attitudes and behaviors I have discussed. It is involved in our everyday interactions, but this is arguably because of deep-seated allegiances, strong desires and any number of other factors – not because the aim is essential to these interactions. To try, in the face of this fact, to justify punishment by appealing to these interactions begs the question against Abolitionism. Begging the question against such an unpopular position is, of course, very tempting. But it is no more acceptable.



### *Conclusion*

The reactive attitudes and their related behaviors do not commit us to punishment. There is nothing essential to them that gives us reason to think punishment justified. Plausible alternative accounts of these attitudes and their associated behaviors are available that are consistent with Abolitionism, the view that punishment is unjustified. The failure to recognize this lies in an inadequate view of these attitudes and behaviors, a simplistic view of the role suffering and harm play in them, and in an inadequate view of punishment itself. I have offered accounts of these attitudes and behaviors that show that the harm and suffering involved in them need not be of a character that furnishes evidence for the justifiability of punishment. The harm and suffering may be important, but not for reasons that support punishment. Moreover, they need not be the product of an aim to impose harm or suffering. Such an aim is essential to punishment. This partial conception of punishment highlights the fact that there are substantive alternatives to punishment. So understood, punishment garners no support from the reactive attitudes and their related behaviors.<sup>11</sup>

Department of Philosophy  
Lawrence University

#### NOTES

<sup>1</sup> I have simplified the formulation of the abolitionist thesis here. Typically, I would formulate it as follows: Abolitionism is the view that the institution or practice of legal punishment is unjustified. The differences between the simplified and more complex formulation are not of much importance here.

<sup>2</sup> While this paper was under review, Bennett also published a book, which elaborates upon the arguments in his paper (Bennett, 2008). I will not explicitly discuss the arguments in his book. The criticisms presented here apply *mutatis mutandis* to them.

<sup>3</sup> I will use the terms *offender* and *wrongdoer* interchangeably.

<sup>4</sup> For more on the consistency of my conception with the expressive conception and the latter's need for supplementation by the former see Hanna, 2008, pp. 124–28.

<sup>5</sup> This is an obvious but unproblematic oversimplification. Whether and how much a mentally ill person can be faulted for his actions is complicated and depends on various factors, e.g. the severity of the illness and its influence on his actions.

<sup>6</sup> There are objections one could still raise to this claim. Perhaps the aim to impose suffering or harm has some sort of conventional status that confers unique expressive powers upon it. Or perhaps the aim is just intrinsically fitting somehow. I respond to these sorts of objections and argue for the claim in more detail elsewhere (Hanna, 2008).

<sup>7</sup> This – namely, the aim's influence on enforcement – is one reason why I take it to be so significant. I am not implicitly appealing to the doctrine of double effect. For more on the influence I think this aim has on enforcement and why I take this to be so significant see Hanna, ms.

<sup>8</sup> For simplicity, I am simply going to talk about beliefs in what follows. It would be more accurate to say we want wrongdoers to acquire various desires and attitudes as well, however. These include negative attitudes towards wrongness and desires not to do wrong. Many of the things I say about the beliefs at issue can also be said about these attitudes and desires.

<sup>9</sup> An anonymous referee offers the following suggestion. Perhaps understanding the wrong one has done or believing that one did wrong just is to suffer (again, I will speak in terms of belief, but many things I say can also apply *mutatis mutandis* to relevant attitudes and desires). The referee suggests that this claim may entail the claim that one cannot aim to generate these beliefs without aiming to impose suffering or harm. I doubt I can do full justice to this proposal here, but I will offer some skeptical remarks. The proposed claim seems ambiguous and the proposed argument seems to trade on an equivocation. The claim can be interpreted as an identity claim, i.e. as saying that the beliefs just are a form of suffering. Or it can be interpreted as a claim about effects, say that these beliefs dispose us to suffer (another referee offers precisely this proposal). I will grant for the sake of argument that the identity reading entails the proposed difficulty about aims (though I actually think it does not). This reading seems false, however. Beliefs – even the sorts of beliefs at issue – are not forms of suffering. This seems intuitively obvious. Furthermore, these states seem to differ with respect to any number of properties. Moreover, there does not even seem to be any necessary connection between these beliefs and suffering. One can imagine someone being pleased and not at all pained by believing (and really understanding) that he has done wrong. So consider the other interpretation about effects and dispositions. It may be true. Even if it is true, however, it does not entail the problematic result about aims. Because these beliefs and the suffering they tend to cause are distinct, one can aim at generating the former without aiming at generating the latter even if there is often a causal connection between them. Given these considerations, I think the referees' proposals can be taken as a challenge to proponents of the arguments under discussion. Can an unambiguous, plausible, defensible claim about the relation between suffering and these beliefs be formulated that entails the denial of my claim about aims? Given some of the issues I have raised, it looks like this project would be fairly complicated and controversial. I think we are justified in being quite skeptical here.

<sup>10</sup> I have said that the harm and suffering involved in guilt and making amends may be bad but that guilt and making amends may nevertheless be good because of some of their other features. With respect to guilt, I suggested that certain beliefs and attitudes involved in guilt might be good instead. With respect to making amends, I suggested the same about the willingness to make amends despite the harm and suffering one might incur. An anonymous referee suggests that these claims entail the following. It would be better if wrongdoers could have all these important beliefs and attitudes about their wrongs without being pained. And it would be better if wrongdoers could make amends without incurring any harm or suffering upon themselves. But, the referee suggests, with respect to guilt, the important beliefs and attitudes are, among other things, dispositions to be pained by one's wrongdoing. And, with respect to making amends, it is doubtful that one could make amends for certain wrongs by means of easy, non-burdensome actions. Even if these claims are true, they pose no obvious problem for my view. Suppose my view does entail that it would be better if amends could be made and certain important attitudes and beliefs held without any resultant harm or suffering. This claim is consistent with the referee's proposed claims about the nature of the attitudes and beliefs involved and about the (im)possibility of easily making amends. The latter do not entail the denial of the former. At best, the latter entail that these states of affairs are impossible. But my view does not entail that these states of affairs are possible, only (perhaps) that they would be better. Now, one might deny that these states of affairs actually would be better, but I do not have space to consider arguments for this view here. A

discussion of Retributivism I offer elsewhere criticizes some of the possible arguments (cf. Hanna, ms.).

<sup>11</sup> Thanks to two anonymous referees for *Pacific Philosophical Quarterly*.

#### REFERENCES

- Bennett, C. (2002). 'The Varieties of Retributive Experience,' *The Philosophical Quarterly* 52, pp. 145–63.
- Bennett, C. (2008). *The Apology Ritual: A Philosophical Theory of Punishment*. New York: Cambridge University Press.
- Duff, R. A. (1992). 'Alternatives to Punishment – or Alternative Punishments?' in W. Cragg (ed.) *Retributivism and Its Critics*. Stuttgart: Franz Steiner.
- Duff, R. A. (2001). *Punishment, Communication and Community*. New York: Oxford University Press.
- 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, *Doing and Deserving*. Princeton, NJ: Princeton University Press, 1970.
- Feinberg, J. (1965). 'The Expressive Function of Punishment,' *The Monist* 49, pp. 397–423.
- Hanna, N. (2008). 'Say What? A Critique of Expressive Retributivism,' *Law and Philosophy* 27, pp. 123–50.
- Hanna, N. (forthcoming). 'Liberalism and the General Justifiability of Punishment,' *Philosophical Studies*.
- Hanna, N. (ms.). 'Rethinking Retributivist Thought Experiments: An Abolitionist Critique.'
- Strawson, P. F. (1974). 'Freedom and Resentment,' in P. F. Strawson (ed.) *Freedom and Resentment*. London: Methuen.
- Ten, C. L. (1987). *Crime, Guilt, and Punishment*. Oxford: Clarendon Press.