

The Retributive Emotions: Passions and Pains of Punishment

ABSTRACT:

It is not usually morally permissible to desire the suffering of another person, or to act so as to satisfy this desire; that is, to act with the aim of bringing about suffering. If the retributive emotions, and the retributive responses of which they are a part, morally permitted or even required, we will need to see what is distinctive about them. One line of argument in this paper is for the conclusion that a retributive desire for the suffering of the wrong-doer, and the aim to bring this about, can (contra recent arguments from Hanna (2008)) be morally justified.

It has been suggested that by reflecting on the role of the retributive emotions in interpersonal relationships, and the alleged legitimacy of the aim for the suffering of the wrong-doer within them, support can be garnered for retributive practices of punishment by the state (Duff (1986, 2001), Bennett (2002, 2003)). The conclusion of the second line of argument in the paper is that whilst the retributive responses can permissibly aim at suffering, the way in which this is so in interpersonal relationships cannot provide support for retributive state punishment.

Responses to wrong-doing, along with resentment, anger, and moral hatred, may include a desire for the suffering of the wrong-doer.¹ Whether such a desire can be justified is one topic of this paper; another is whether the role such a desire plays in interpersonal responses to wrong can provide any support for retributive punishment by the state.

Why focus on this aspect of the retributive emotions? First, it is not usually morally permissible to desire the suffering of another person, or to act so as to satisfy this desire. Nonetheless, some retributivists have recently argued that in interpersonal relationships, the suffering of the wrong-doer is appropriate in response to wrongdoing, and that it is permissible - perhaps even required - to try to bring such suffering about. I will argue in support of a particular understanding of these claims (sections 1-3). This first part of my argument responds to a recent claim (from Hanna 2009) about the legitimate aims of the retributive responses.

Second, it has been suggested that the role of the retributive emotions in interpersonal relations provides support for retributive practices of punishment by the state (see Duff, 1986, 2001; Bennett, 2002, 2003). For example, Duff writes that 'blame, like punishment, is intended to induce suffering in an alleged offender for an alleged offence', and goes on to remark, with regard to criminal punishment, that 'an essential part of its meaning and justification lies in its relation to moral blame' (1986, p.41). I will argue against this strategy: neither the suffering of guilt, nor the suffering of repentance, as they feature in the 'retributivism of interpersonal relations' can provide support for the imposition of suffering of punishment (sections 3-4). This part of the argument calls into question a methodology prominent in contemporary literature; namely, to draw support from interpersonal analogues for state practices of punishment (see Bennett 2002, 2003; Duff 1986, 2001; Tasioulas 2006).

1 See Strawson 1974; Bennett, 2002, 2003; Mason 2003; and Murphy & Hampton 1990.

1. Retributive responses and aiming at suffering

Imagine that you arrive home to find the lock broken, the door hanging, your belongings ransacked: a burglary. You feel a range of emotions: anger at the loss; indignation at the disrespect manifested in the break in; resentment towards whomever it was that disregarded your interests, and the kinds of inconveniences the damage and losses will now cause you. Perhaps you also feel also dismay that such considerations were paid little heed. You might also want to see the wrong-doer suffer – to feel bad about the wrong; or to pay for what he or she has done.

Is the desire for suffering a necessary component of the retributive emotions? Strawson (1974) remarks on:

'the partial withdrawal of goodwill which these attitudes *entail*, [and] the modification they *entail* of the general demand that another should, if possible, be spared suffering' (1974, p.23, emphasis added, original emphasis removed).

The reactive attitudes of anger and resentment entail a withdrawal from the usual unwillingness to impose (or see) suffering. Further:

'the readiness to acquiesce in the infliction of suffering on the offender which is an essential part of punishment *is all of a piece* with this whole range of reactive attitudes' (p.23, emphasis added)

On this characterisation the desire for the suffering of the wrong-doer (the readiness to acquiesce in its infliction) is a part of the usual range of the reactive attitudes of resentment and anger. Strawson casts such a desire as a common component of the reactive attitudes, along with resentment and anger.

One might maintain that far from being necessary, such desires have no proper place in the reactive attitudes to wrong-doing. Darwall for example, maintains that 'reactive attitudes like resentment and indignation can be distinguished from the desire to retaliate or gain vengeance' (2006, p.83). On his view, a vengeful wish for suffering responds to disrespect with disrespect; though resentment and indignation presuppose respect for the moral authority of the person blamed or resented (p.84).

Darwall's suggestion presents an appealing ideal.² However, I take Strawson's claim as reasonably uncontroversial: the desire for the wrong-doer's suffering is at least sometimes - perhaps frequently - a feature of human responses to wrongs. The question is whether, when it is present, such a component of the retributive responses can be justified. Are such desires rightly understood as a form of disrespectful vengeance? Do some desires for suffering have a proper place in the emotional responses to wrongdoing?

I will argue that, on a suitably narrow understanding, such desires can be justified. As such, I continue to refer to such responses as 'the retributive emotions', but it should be noted that the retributive element I argue is warranted is narrower than that for which others have argued.

1.1 Retributive hatred and the exclusivity requirement

In the face of concerns that the retributive responses are 'vindictive, inhumane, barbarous and immoral'³ theorists have tried to rationalise responses, such as retributive hatred, arguing that it can be morally permissible, and perhaps even required, to desire the suffering of someone who has wronged you. Such responses, it is argued, are in the service of some value; they are not base and brutish. Consider Murphy's remarks:

² For discussion of empirical work that may provide support for Darwall's picture as an ideal, see Holroyd 2007.

³ Mabbott, quoted in Armstrong, 1961, p.471. See also Bennett 2002 p.148.

‘Jones has injured me, has taken unfair advantage of me, has brought me low, and is himself unrepentant and flourishing. I hate him and want him brought low. [...] I want Jones to be hurt. [...] [P]art of the basis for desiring the hurt is the desire to restore what seems (at least to me) to be the proper moral balance of whatever goods are in question’ (1990, p.89)⁴

Such hatred, and the wish to see another suffer, he claims, is neither irrational nor immoral (p.90); rather, it is the appropriate response and serves to maintain one's self-respect: 'resentment, (in its range from righteous anger and righteous hatred) functions primarily in defence of ... certain *values of the self*. [...] [T]he primary value defended by the passion of resentment is *self-respect*' (p.16). To recognise that one has been wronged, undeservedly so, requires a response: self-respect requires that one feel resentment when wronged, and seek retribution and reparation for that wrong. To fail to do so is to allow oneself to be slighted, to be a 'doormat'.

Murphy suggests that various considerations might ultimately counsel against acting on such desires,⁵ but he nonetheless maintains that they are 'natural, fitting and proper [in] response to certain instances of wrongdoing' (p.108). In virtue of what is the response 'proper'? Whilst it seems quite right that self-respect requires some response, there is an argumentative gap between the claim that self-respect requires that wrongs do not go unchallenged and the claim that self-respect requires that one ought to desire (and perhaps bring it about) that the wrong-doer suffer.⁶ Other responses that preserve self-respect might yet be available – for example, one might desire that the offender recognise that he has wronged you, or that the community publicly reaffirm your rights.

Hampton, for example, offers an alternative which does not make suffering a necessary part of the retributive response:

'punishment is not so much the infliction of pain as it is the infliction of a *defeat* that annuls the wrongdoer's claim of superiority, which such non-painful experiences as community service can deliver, just as well as pain (although because pain is commonly seen to represent defeat, it is a useful medium to symbolise the idea that the wrongdoer is not one's superior)' (1990, p.143)

Although Hampton is similarly concerned with the assertion of one's moral status in response to wrongs which attempt to degrade it, she envisages responses that affirm self-respect and are non-painful.⁷ Pain might be a useful symbol, but alternatives are available. Reflecting on this argumentative gap indicates what I want to draw out as a moral constraint on our responses to wrongdoing; namely that where such responses are guided by some purpose, the method of securing that purpose which involves the least suffering should be preferred. More precisely:

4 Murphy 1990.

5 To wit: we might not be in an epistemic position to know that our hatred is warranted; it may be costly or impossible to get even; we may be in no position to make such judgements; or may find that moral decency counsels against acting on retributive desires.

6 Murphy does try to fill this gap. In brief, relying on a principle according to which it is morally permissible to *desire to do* (under the same description) what it is morally permissible *to do*, he maintains that it is morally permissible to punish (that is, to impose such suffering); and therefore, it is morally permissible to desire to do so (p.94). First, I doubt the principle is generally true. Certain 'lifeboat ethics' actions, (such as that of 'saving the five at cost of letting one die') or the imposition of certain risks (e.g. risks to life imposed by cyclists) may be morally permissible. But to desire to do such things appears morally problematic. Such an actions are, if morally permissible, regrettable. Desiring to so act would be morally problematic. Second, Murphy's argument relies on the assumption that punishment is permissible, and I am not yet willing to grant this. Finally, Murphy himself concedes that there may be a number of reasons for overcoming the responses of retributive hatred (see fn.6).

7 One might wonder whether options such as community service do not involve suffering (see Duff 2001, pp.99-104). See section 4.

Exclusivity: if it is permissible to do A, which imposes suffering for purpose P (e.g. affirming self-respect, expressing censure, communicating disapproval and making reparation⁸), this is because no (feasible) alternative way of securing P which imposes less suffering is available.

This seems to be a plausible and morally respectable general principle, and one which can guide thinking about the retributive responses. Those retributivists who claim that the imposition of suffering is necessary for some further purpose will need to show why imposing suffering (to that extent) is the uniquely appropriate way of securing or protecting this further goal or value (either that it is instrumentally superior, or that it is a constitutive part of that goal). In particular, these retributivists must show the inadequacy of alternative responses to wrongdoing (such as Hampton's) which do not require that suffering is imposed. Thus the burden is the heavy and ongoing one of arguing for the superiority of imposing suffering over other responses which do not require this.

On some versions of retributivism the imposition of suffering in response to wrongdoing is an intrinsic good.⁹ Such views will not be susceptible to the concern that the valued purpose can be secured by means which do not involve (or at least minimise) suffering; for proponents of such a view hold that the suffering in itself is the valuable aim. My argument will not speak directly to such views. That much of the recent literature attempts to move away from such views, indicates, in my view, their implausibility. I hope that by showing the difficulty of justifying the imposition of suffering even for the sake of some plausible aim (censure or affirmation of self-respect, say), the difficulty of arguing that it is intrinsically valuable becomes clearer.¹⁰

2. Expressions of disapproval and the exclusivity requirement

Recent explorations of the retributive responses provide a clear view on the ways (if any) that the retributive responses might legitimately have the suffering of the wrong-doer as an aim. Bennett has prised apart the painful elements of the interactions that follow wrong-doing in interpersonal relationships: the expression of disapproval, the acceptance of blame, and the burdens of repentance and reparation. These aspects are painful for the wrong-doer, but engaging in them is not a matter of mere retributive hatred; they are constructive parts of calling the wrong-doer to account for her action.¹¹ I deal here with the first aspect of retributive suffering – expression of disapproval – turning to the latter two in following sections.

We have not seen that it is appropriate to desire or aim to impose suffering upon the wrong-doer, when wronged. But it is surely plausible that it is permissible, and even required, to express disapproval in response to wrongs. To not do so really would fail to adequately address the wrong, and perhaps fail to manifest the appropriate self-respect. Recall the example of the burglary, above: encountering the burglar (later), it would be quite proper – and certainly morally permissible – to call her to account ('Why did you do it?', 'Didn't you think of the upset it would cause?') and to express disapproval.

But expressing disapproval, Bennett argues, itself involves imposing suffering. It involves behaviours (by the wronged or perhaps some third party) which are typically painful for the recipient: withdrawal from a relationship, is Bennett's paradigm case. By refraining from speaking to the wrong-doer, or 'giving them the cold shoulder', we convey disapproval and impose unpleasant

⁸ See Murphy 1990; Tasioulas 2006; Duff 2001 respectively.

⁹ E.g.: 'there is some intrinsic value in the suffering of the guilty' (Davis, 1972, p.136.); this is 'seen as right or good in itself, apart from the further consequences to which it might lead' (Nozick, 1981, p.374)

¹⁰ Such views face difficulties in explaining why suffering other than that imposed by the state (e.g. suffering which precedes conviction) is not relevant to how much punishment is deserved. See Metz 2000, pp.506-507.

¹¹ See also Duff 1986, p.42-52; 2001, pp.107-112.

treatment upon her. Similarly, Duff writes of 'the pain which I suffer merely from the hostile reactions of others' (1986, p.59). These reactions are painful to experience.¹² But such treatment is not simply retributive hatred; its aim is not to 'bring the offender low'; rather such treatment is an expression of moral disapproval which attempts to bring it about that the offender recognises the wrong done: such withdrawal is expressive and purposive (Bennett 2002, p.152).

Is it justifiable to impose such suffering? This will depend upon whether suffering is a necessary part of the communication of disapproval. If so, then, insofar as it is justifiable to communicate disapproval, imposing suffering as part of that communication can be justified. If the retributive emotions are comprised of a desire for *this* kind of suffering - suffering due to the unpleasant expressions of disapproval - then such a desire may be justified. But if such suffering is not necessary to the communication, then the exclusivity condition will not be met: an alternative way of communicating without imposing (as much) suffering is morally preferable.

2.1 Withdrawal as a necessary means of expression?

There are two reasons to suppose that Bennett does not pick out a necessary component of the retributive response. First, there is reason to doubt that withdrawal is necessary for the expression of disapproval. Second, alternative means by which disapproval may be expressed need not be painful for the wrong-doer. I deal with these issues in turn.

Bennett uses an example different from mine; he focuses on relationships already established (between colleagues), from which some parties, in response to the wrong, withdraw.¹³ Thus, he claims 'retribution *requires* the suffering that is imposed on the wrongdoer by the withdrawal of other members of the community' (2002, p.152 (my emphasis)). Is withdrawal the only, or even best, way to communicate disapproval?

My example (of the burglar) suggests that claims about the role of withdrawal in retribution are limited. For my example provides a case in which (we have supposed) no pre-existing relationship obtains, and thus there is not *withdrawal*, as such, from the relationship. Rather the relationship is established by the wrong, and by the responses to it. It cannot be right that withdrawal is *required* in all cases for the communication of disapproval. Bennett's claims, if they are true of some relationships, cannot be extended to responses to wrongdoing by strangers. This leaves open the possibility that means other than withdrawal are available to communicate disapproval.

Might one insist that nonetheless there is a kind of withdrawal in my example? We could say that one ceases to have the dispositions one might otherwise have towards the offender; you are no longer well-disposed towards her, as you are to others - there is no longer a presumption of goodwill between you. Withdrawal from *this* stance, if not from pre-established relationships, is painful for the offender. But casting the withdrawal claim in this way seems forced: first, because the absence of prior interaction between the two parties; second, because the wrong committed has already broken that presumption; and third, because given the sparseness of our existing relationship, there is a limit to how much I can withdraw without failing in my basic duties. I resent the wrong-doer, but I would not (and ought not) fail to react if it were she drowning in the nearby pond.

If painful withdrawal from relationships is not an essential component of the retributive responses, what alternative modes of conveying disapproval are there? Must conveying disapproval be

12 Morris (1971) also refers to 'the pain that comes from separating ourselves from the union that we value' (p.426), i.e. from the relationships before the wrong. This metaphor is a little opaque, so I focus on Bennett's claims - although my concerns (in 2.1) apply to this aspect of Morris' claims also.

13 One might argue that it is only appropriate to focus on such close relationships in considering interpersonal responses to wrong; but this seems to be a mistake. Many wrongs happen between individuals who do not know each other well (if at all), and the reactive attitudes are a feature of these 'minimally' interpersonal relationships.

painful? Or can one communicate blame and disapproval whilst neither wanting, nor aiming for, the infliction of suffering?

Consider Hanna's critique of Bennett's view. He asks:

“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. [...] We can, for example, express disapproval in careful, measured ways and even offer comfort to those we criticize” (2009, p.241)

On Hanna's view aiming to express disapproval does not entail that one aims at the suffering of the wrong-doer. Other expressive means - albeit less conventionally familiar ones - could be used to convey moral disapproval. So on this view, in accordance with the exclusivity condition, if one aims to express disapproval one ought to do so in the way that involves the least suffering.

2.2 Expressing disapproval and imposing suffering

Hanna's remarks here are ambiguous. On one reading, Hanna is asking whether the *means of expression* must aim at suffering; that is, whether the expression of disapproval must use a means of conveying disapproval which is intentionally painful for the recipient. Withdrawal from a relationship (imposing psychological harm), or corporal punishments (imposing physical harm) are examples of painful modes of expression. This reading is implied by the answer Hanna gives to his own question, above. The other reading, which I take up in more detail shortly, concerns whether *expressing disapproval itself* must aim at suffering.

Hanna's claim that the *mode of expression* of disapproval need not be one which involves suffering is plausible. As we have seen, painful withdrawal from relationships is not the uniquely appropriate means of expressing disapproval. Other ways of expressing this are available: simply telling another, or writing to them. In accordance with the exclusivity condition, they are morally preferable to modes which impose suffering.

But are such alternative modes of expression really adequate to convey disapproval? One might argue that such alternatives are in fact *not* sufficient to communicate moral disapproval in all cases; sometimes, harsh treatment is needed to express blame. For example, Tasioulas maintains that to properly convey moral condemnation of some serious wrongs, hard treatment is required:

'only punishment adequately conveys the blame the wrong-doer deserves. [...] *Only* such a response adequately reflects the gravity of the wrong that has been committed. [...] It] enables us to properly evince the sincerity of our condemnation' (2003, p.296, emphasis added).

Is this right? One can make various utterances to communicate blame, and the meaning of such utterances is clear. But the thought here is that a response to some serious wrongs which did not impose suffering would be inadequate to show that one *really* disapproves or blames the wrongdoer; it would not be enough to show that one takes the wrong seriously. We can agree that it is important to respond in a way that takes the wrong seriously. But must this involve the imposition of suffering? Tasioulas remarks that we should see that it is appropriate as 'a basic norm of justice' (2006, p.297). But it is important to consider whether there are other responses that are able to convey blame whilst also taking the wrong seriously. Whilst further work is needed to show that other such responses could be adequate, I here want to show that Tasioulas' argument for the claim that only punishment is adequate is based on a false assumption about the relationship between blame and punishment.

Central to his argument is the claim that that only by imposing suffering (e.g. by state punishment) can one communicate a condemnation that is keyed to the gravity of the wrong. Tasioulas is able to move from the plausible claim about the appropriateness of blame or condemnation to the contentious claim about appropriateness of imposing painful punishment as a means of conveying blame by 'rejecting the stark distinction between censure and hard treatment' (2006, p.295). On his view, the imposition of punitive suffering is just a 'scaled up' version of the condemnation of blame (scaled up in proportion to the seriousness of the wrong).

However I will shortly argue (in section 3) that it is precisely right that there is a stark distinction between moral disapproval and the hard treatment of punishment - that they are importantly different in kind. In particular, expressions of blame ought not impose suffering. So the smooth transition from the appropriateness of blame to the appropriateness of imposing punitive suffering cannot be made.

State punishment may have some distinct functions, such as the 'authoritative disavowal' of the wrong (Feinberg 1965): punitive responses may serve to distance the state from the wrong (Metz 2000, p.497). One might argue that the imposition of suffering is needed to perform this function and to convey condemnation to *others* (this is not usually important in interpersonal relationships, and may be inappropriate). To write a letter, or even make a public announcement in condemnation of the wrong, would be insufficient to symbolically distance the state from the wrongs.¹⁴ Tasioulas's remarks about evincing the sincerity of the condemnation may draw on this thought.

Even still, a further argument is needed to show that suffering is the only appropriate way for the state to respond adequately to a wrong: there may be alternatives that are non-painful but can still demonstrate that one takes the wrong seriously. Indeed, a response which is constructive and likely to help the wrong-doer avoid further instances of wrong might do more to demonstrate a sincere disavowal of the wrong.¹⁵ More work is needed to show the adequacy of such alternatives. Importantly, any such response will be coercively imposed. As I will argue in section 4 this is another reason to suppose that such practices cannot garner support from our responses to wrongdoing in interpersonal relationships.

To recap the claims of this subsection: Hanna's remarks challenged the thought that the *means of expressing* disapproval must impose suffering, and I have argued in support of Hanna's claims, against those who maintain that only by imposing suffering can adequate blame or disapproval be conveyed. Some of this argument, regarding the distinction between expressing blame and imposing hard treatment, is still pending. We can now turn to this matter, in considering the other possible disambiguation of Hanna's remarks.

Recall the alternative reading of Hanna's claims: must *expressing disapproval itself* aim at suffering – a suffering independent of the means of expression used? On this interpretation, the issue concerns whether expressing disapproval, even by moderate means such as telling another, or writing to them, entails aiming at suffering. Do we suppose that being on the end of moral disapproval, no matter how it is communicated, is painful and unpleasant? If we want someone to know that we disapprove of them, does this entail that we desire that they undergo some suffering - and do the retributive responses thereby aim at suffering in this sense? Hanna's disambiguation of his question did not address this cluster of questions. But they are of central importance, as we shall see.

14 In this paper Metz is explicit in assuming, rather than arguing for, the claim that harsh treatment can be justified as the means of expressing censure.

15 An example is the 'Changing Lives Through Literature' programme, which offers mandatory reading groups as an alternative sentencing option. See <http://ctl.umassd.edu/home-flash.cfm> (accessed 12/08/2010).

3. Expressing disapproval and inducing suffering: guilt

In this section, I make three key claims: first, that the kind of suffering that is justified in response to wrongdoing is the suffering of guilt; second, that it is permissible to act to bring about this suffering; third, in so acting it is permissible to act with the aim, or desire, of bringing about this kind of suffering.

I will start with part of the justification for the second claim. We have seen (in section 2) that, plausibly, expressing disapproval in response to wrongdoing is permissible (and might even be required by self-respect). Additional support for this claim is found in the purposive element of conveying disapproval, to which Bennett (2002) has drawn attention. He maintains that, when purposive in this way, the retributive emotions do not merely issue in cathartic behaviour:

it would be wrong to understand the expression of the emotion of blame as a mere release of pent-up emotion, such as might be taken out on a punch-bag (p.151)

Rather, the aim is to bring about a certain change in the wrong-doer:

The behaviour which expresses blame is, after all, an attempt to bring something about, namely, that the offender should recognize that we disapprove of him (p.153).

Not only do we try to communicate the fact of our disapproval; we aim to bring the wrongdoer to share it (p.152)

The aim of expressing disapproval is to bring it about that the wrong-doer recognises that she has done wrong, and accept that disapproval is warranted. Thus additional justification for blame is grounded in the goal of bringing the wrong-doer to acknowledge the blame; this may be an important step in their coming to repudiate such actions.¹⁶ Having done so, we expect the wrong-doer to resolve to refrain from so acting in future; she might reaffirm the values that she now recognises were violated, and shape her future behaviour accordingly.

But (moving now to the first key claim) in bringing it about that the wrong-doer accepts disapproval, or blame, for the wrong, Bennett claims, we also bring it about that the wrong-doer experiences a certain kind of suffering, for:

when [s]he comes to see ... that [s]he has done wrong ... there is the painful feeling of guilt. Guilt is a kind of pain you impose on yourself (2002, p.155).

It is painful to know that one is prone to fall short of [moral standards] or more seriously, to harm the things one aspires to respect. So feeling guilt follows on in some sense (not automatically, but in terms of a deepening moral understanding) from an understanding that one has done wrong (2003, p.134).

The kind of suffering at issue here is the unpleasant pain of guilt. As Herbert Morris puts it:

'reflecting on the ways we have felt when feeling guilty, we may recall thinking that we felt rotten, depleted of energy, and tense' (1971, p.246).

This kind of suffering seems an appropriate, and indeed constructive, part of the retributive experience. It is what Duff refers to as 'the kind of pain which is the proper end of moral blame' (1986, p.59). If this is the kind of suffering at issue, and if it is permissible to act (by expressing

¹⁶ There is a difference between getting the agent to realise that her action was wrong, and getting her to realise it is disapproved of. The former is principally at stake here. See Duff 1986, see pp.42-52.

disapproval or blame) to bring it about, should we also say, then, that the desire or aim to bring about this kind of suffering is permissible?

One might think not: I can act with the recognition that guilt is an appropriate reaction (as the 'proper end' of blame), without it being permissible to desire that such a reaction comes about. By analogy, I might act with the recognition that grief is an appropriate reaction to my imparting the news of a death, but whilst I anticipate the recipient will feel grief, surely I ought not wish for, or desire that she feel it? Indeed Hanna has recently rejected the claim that aiming at suffering is a necessary part of expressing disapproval. He maintains that:

‘aiming to get offenders to adopt such beliefs [about their past wrongdoing and the appropriateness of blame] need not involve an aim to harm them or make them suffer’ (2009, p.245).

In support of this claim, Hanna argues that there are two ways to understand the claim that to aim for the wrong-doer's acceptance of blame is to aim for their suffering. Neither, he claims, are plausible. Thus even if acting in a way that brings about such suffering is permissible, aiming to bring it about may not be. What are the two interpretations at issue?

3.1 Inducing guilt and aiming at suffering

First, the claim could be an identity claim: to have the relevant belief *is* to suffer. What looks very much like an identity claim can be found in Duff's remarks:

‘I cannot first blame [the wrong-doer] and then decide whether to try to make him suffer thus, since to blame him *is* to try to induce such suffering; he cannot accept my blame and hope to avoid such suffering, since to accept blame *is* to suffer thus’ (1986, p.60 emphasis in original)

But this, Hanna claims, is implausible: a belief is not suffering. Moreover, one can imagine counter-examples to the identity thesis: someone who accepts that he did wrong and is gleeful at this thought.¹⁷

On the second understanding, there is a causal connection between accepting the relevant belief and undergoing suffering: ‘these beliefs dispose us to suffer’ (Hanna, 2008, p.249 n.9). Many of us may as a matter of empirical fact feel bad in response to accepting blame. But we can certainly think of cases where this causal relation does not hold. Even in the cases where the causal relation does obtain, Hanna suggests, it does not follow that those expressing disapproval aim at the suffering of the wrong-doer, nor that they ought to.¹⁸

It seems quite right that neither of these two understandings of the relationship between accepting blame and the relevant self-imposed suffering can generate the conclusion that those expressing disapproval permissibly aim at suffering. But Hanna has offered us only two options, when there is a third.

An alternative option (and the sense in which we should read Duff's claims above) is that when one comes to believe that one has done wrong and accepts blame, one *ought* then to feel bad. Proper acceptance of blame is in part *constituted* by such feelings of guilt. So the claim about the belief and the suffering is neither an identity claim, nor a causal claim, but a normative one. That the suffering

¹⁷ Some may deny that the person who responds in this way really does believe (or judge) that she did wrong: she may use moral notions only in the ‘inverted commas’ sense (Hare 1952, pp.124-126).

¹⁸ I presume the thought here is: I can aim to win the prize without aiming to make the loser cry, even if I believe it will most likely occur as a causal consequence of my winning. And, in the usual course of things, I certainly ought not to aim for this.

of guilt is a constitutive part of accepting blame seems to be what Bennett and Duff have in mind in their respective comments: that the interaction is 'fundamentally incomplete unless it brings about guilt and repentance' (Bennett 2002, p.155); and such 'pain is integral to a repentant recognition of guilt' (Duff 1986, pp.59).

This third option is the most plausible one. It assumes that affective states are constitutive of certain moral attitudes towards one's actions: fully recognising that one has done wrong is constituted by the pain of moral guilt.¹⁹ The claim is not a causal one: an agent may fail to feel bad, in the way that they ought to. But this will indicate a, rational failure, or a failure to properly recognise the wrong and accept the moral disapproval.²⁰ If this is right, then aiming to bring about in the wrong-doer acceptance of blame cannot be prised apart from aiming to bring it about that she feels the suffering of guilt. Hanna is wrong to suppose that accepting blame is simply a matter of forming beliefs: the suffering of guilt is constitutive of the proper acceptance of blame. Expressing disapproval with the aim of bringing the wrongdoer to share in that disapproval cannot be separated from the aim of bringing about the suffering of guilt.

If these arguments are accepted, then we have secured the conclusion that it can be permissible, in this way, to desire or aim for the suffering of another. Whilst such a desire is not usually morally permissible, consider the various further things that can be said in justificatory support of desiring and aiming for the suffering of guilt: first, if this suffering is constitutive of the acceptance of blame, the exclusivity condition is met. The relevant purposes of the interactions following wrong-doing (moral repair and reconciliation, via the expression and acceptance of moral disapproval) cannot be achieved without this kind of suffering.

Second, having this aim is not susceptible to the usual objection that a desire for suffering is vindictive and barbaric. The suffering involved here - the suffering of guilt - is a constructive part of the response to wrongdoing. Moreover, practices of blaming treat the wrongdoer with respect, as a person capable of change. As Tasioulas emphasises, 'implicit in the blaming response is a recognition of their moral status as responsible agents beings who ... are able to grasp and comply with reasons' (2006, p.294). Finally, if the pain of guilt is constitutive of the proper recognition of wrong and acceptance of blame, then considerations of self-respect make it reasonable to seek such a response.

I have argued that following wrong-doing it can be permissible to desire, and act with the aim of inducing, the pain of guilt. This suffering is constitutive, and normatively required, of those who properly accept blame for their wrongdoing, and there are good reasons for desiring that the suffering of guilt are felt, and for aiming to bring it about. But can this preliminary conclusion, lend any support for the practices of punishment?

3.2 Justifying punishment?

Hanna argued that because responses to wrong-doing need not and should not aim at suffering, they do not provide support for punitive institutions. I have argued that our retributive responses can permissibly aim for the suffering of the wrong-doer, in a specific sense. Does this help to justify retributive punishment by the state? My aim in this subsection is to show that the 'retributivism of interpersonal relationships' cannot provide justificatory support for retributive state punishment

19 Compare judgement internalism about moral beliefs, where judging an action morally good or right entails a (defeasible) motivational pull. See Brink, 1989 p.40; Darwall, 1983, p.54.

20 One might maintain that we should still hope that suffering is avoided. Might the hope be that agent experiences moral or rational shortcomings, such that she does not feel bad as she would otherwise? This strategy is implausible: we can hardly call for the failure of moral and rational capacities so as to avoid suffering, if those very capacities ought to be engaged when we ask someone to accept blame. Moreover, it is unclear that reconciliatory processes could be achieved in the absence of such guilt: part of the process of accepting disapproval and undertaking to reform requires that one feel bad in the appropriate way.

(hereafter, simply 'punishment').

According to retributivists, punishment ought to aim at suffering, either for its own sake (on the 'intrinsic good' versions of retributivism); or for the sake of some further purpose (instrumentally, or as constitutive of), for example, the expression of censure, or the communication of disapproval. For these ends, it is held that the appropriate (and sometimes only adequate) response to criminal activity is to impose some painful experience upon the wrong-doer. Of the various justifications for such views, I am concerned here solely with whether interpersonal practices of blaming can support retributive state punishment.

In the first two sections, I argued that in interpersonal relations it is difficult to justify aiming at *imposing* suffering. Insofar as disapproval or condemnation can be expressed without the *means* of expression being painful, to aim at imposing suffering cannot be morally permissible. I then argued that it is permissible to aim to *induce* a certain kind of suffering, namely, guilt.

The first thing to note is that the kind of suffering that it is permissible to aim at in interpersonal relationships²¹ and the manner in which it is brought about, is contra Tasioulas (see subsection 2.2), quite different from the kind of suffering aimed at in punishment. Morris makes this clear:

'the pain we suffer when feeling guilty differs from the pain that is inflicted with the significance for us of punishment' (1971, p.129)

And this much is recognised by Duff, in his acknowledgement that whilst the proper purpose of blame is:

'to induce the kind of pain which flows from an understanding of the condemnation ... matters are surely quite different when punishment takes the form of hard treatment: for the suffering which such punishments most obviously and directly cause is independent of their communicative meaning' (1986, p.242).

Any painful guilt felt due to recognition of wrong-doing and acceptance of blame is supplemented, in punishment, with other forms of suffering. In punishment, the forms of treatment imposed (incarceration, fines) are intended to be unpleasant – they are supposed to be burdensome and unpleasant to undergo.²² Any aim to *impose* suffering cannot garner support from the reflections on blaming in interpersonal relationships. For we saw that it is permissible to desire and act so as to *induce* suffering of a particular kind: guilt. Not only is guilt a qualitatively different kind of suffering to that of punishment, it is a form of suffering that is induced, and flows from the wrong-doers own understanding of her action, rather than being imposed by another.

Perhaps, then, punitive practices find support insofar as they aim to induce in offenders guilt? But it is not clear that retributivists envisage guilt as any thing more than a desirable side-effect of punishment. Nor should they, without significant argument to show that bringing about specific affective states (namely, guilt) in its citizens is the legitimate business of the state, and is consistent with respect for values such as autonomy and privacy.

Moreover, if we are primarily concerned with attempts to induce guilt in the wrong-doer, there is evidence to suggest that imposing suffering is not likely to be an effective way of doing so. In brief, social psychologists' findings indicate that individuals tend to react oppositionally to judgemental pressures and attempts at controlling behaviour. An agent who experiences this kind of 'reactance'

21 I mean relationships between people (close friends and strangers) rather than between citizen and state.

22 I return to the matter of treatments that are not intrinsically burdensome, but which are coercively imposed, in the next section.

will then be less likely to take on board the message of the judgement, and more likely to be defensive of her action and the values that informed it (Springer 2008, p.410-412). On the basis of this evidence we would expect the imposition of suffering qua punishment (an acute form of judgemental pressure, and an obvious attempt to control behaviour) to be a singularly *bad* way of bringing about the acceptance of blame, and the suffering of guilt.

The suffering permissibly aimed for in practices of blaming provides no support for, and perhaps provides some considerations that speak against, imposition of punishment in response to wrongs. To be clear: I have not argued that such suffering (or the aim to bring it about) cannot be justified, but rather that any justificatory support cannot be garnered from reflection on the retributive emotions as they feature in our practices of blaming.

4. Repentance, conditionality and coercion

I will briefly consider a further kind of suffering identified as important to interpersonal retributive responses: that involved in repentance. Showing repentance might require making amends, by undertaking burdensome steps which express regret. Repentant acts may be part of an attempt to repair the relationship damaged by the wrong:

A repentant wrong-doer might also punish himself ... [and] undertake and undergo a penance as a punishment, not simply because [he] deserves to suffer for it, but to assist and express [his] repentant understanding of what [he] has done (Duff, 1986, pp.68-69).

Bennett too suggests that:

[A wrong-doer] might be expected to do something more, something which involves some extra sacrifice to himself (2002, p.160. See also 2003, pp.134-136).

Might this further kind of suffering provide some support for retributive state punishment? Consider first the characteristics of the suffering of repentance. It is undertaken in an attempt to repair damaged relationships. Such 'penances' need not be intrinsically unpleasant, but the undertaking of tasks in penance can itself be painful, expressing moral humility, one's acceptance of blame, and willingness to put things right. One important aspect of these sacrifices is that they are *self-imposed* and *willingly* undertaken: the burdens are part of a voluntary undertaking which has symbolic value:

‘The *willingness* to undergo these forms of suffering is a sign of sincere repentance’ (Bennett, 2002, p.161, my emphasis).

Likewise, Duff is clear that the suffering of repentance is not necessary, but can be voluntarily taken up:

‘I do not suggest that anyone who sincerely repents his wrongdoing *must* express this through a penance [...] My claim is only that he may intelligibly feel the need for a penance’ (1986, p.69)

It may be misleading to claim that the suffering of repentance is entirely self-imposed. As a step towards reparation, some (perhaps tacit) negotiation about what sort of burdens or how much suffering could repair the relationship, is likely appropriate. In such cases, the suffering of repentance will be conditional: if relations are to be restored, then some burdensome suffering ought to be undertaken. Precisely how much will depend on the nature of the wrong, and on the nature of the relationship damaged.

These features of repentance are important. Following wrong-doing, one might desire that the

wrong-doer suffers the burdens of repentance. But it would be wrong to aim to impose it: for repentance is voluntarily undertaken and conditional upon the wish to repair relationships. It may be that the relationship cannot be restored, or that it is not something the relevant parties want. Consider Bennett's example: Bryson has cheated on his partner, has accepted blame and feels guilt. Should his partner then seek that Bryson suffers for the sake of repairing the relationship, taking on burdens in order to make up for the wrong? If the relationship is to go on, then it is likely that such steps will be required. But perhaps whilst Bryson is sorry and feels bad about the harm he caused, he does not want to repair the relationship. Or, he may wish to do so, but finds the penitential burdens needed to show his repentance too great: he may not want a relationship in which he is always making amends. He may accept the blame, but decide not undertake a penance, and to leave the relationship.

This example brings out another important feature of this aspect of retributive responses: in the case where it is decided that relationships will not be restored, it would be wholly inappropriate for the wronged party to insist on repentance. It would be wrong, for example, to detain or coerce Bryson into undertaking burdens; or to persist in demanding that reparative steps are taken. Whilst in interpersonal relationships repentance and its 'penitential suffering' can play an important role in restoring relationships that have been damaged by the wrong, such suffering is (perhaps with negotiation) self-imposed, and conditional upon the desire to restore the fractured relationship.

4.1. Justifying punishment?

I argued that the suffering of guilt cannot provide support for retributive state punishment. Can repentance? A number of theorists have suggested that we can see the suffering of repentance as an analogue for the suffering imposed by punishment. Duff argues that punishment is appropriately thought of as a secular penance. He talks of:

'the part which hard treatment can properly play ... (it serves as a penance which the criminal should ideally come to will for himself)' (1986, p.245);

'we should understand and justify criminal punishment ... as a species of secular penance' (2001, p.106).²³

Again I want to briefly suggest that the interpersonal case is not a good analogue to state practices of punishment and cannot provide support for them. First, the kind of suffering involved in reparation is self-imposed and voluntarily undertaken. As such, it differs significantly from the kind of suffering involved in punishment, where it is imposed by some third party.

Duff suggests that the imposition of suffering might be needed to focus the wrong-doers attention on the wrong:

'it provides a structure within which, we hope, he will be able to think about the nature and implications of his crime, face up to it more adequately than he might otherwise (being human) do, and so arrive at a more authentic repentance' (2001, p.108)

But in earlier discussion I mentioned empirical considerations which indicate that imposing

²³ Bennett 2008 has also attempted to connect the interpersonal 'cycle of blame and apology' to state practices of punishment. His account differs in structure from Duff's (on which the hard treatment (amongst other things) represents the penitential suffering of the wrong-doer. For Bennett, hard treatment expresses how wrong we think an act is, symbolising how sorry that person ought to be – that is, how much they would have to do to make up for the wrong (2008, pp.146). I cannot do justice to these claims here, but will for now note that it is not usually appropriate, in interpersonal relationships, to impose suffering in accordance with how sorry we think another ought to be. Earlier concerns also arise, about whether non-painful modes of expression can have appropriate symbolic force. Detailed consideration of Bennett's claims in this regard is reserved for another occasion.

suffering is not likely to encourage reflection on past wrong-doing, so we have reason think Duff's claim here are unduly optimistic. In any case, Duff's remarks here suppose that the wrong-doer has already 'signed up' to repentance: suffering helps him to repent 'authentically'. However, if the individual has not 'signed up' to repentance (he need not, we saw from considering interpersonal relations), this justification cannot apply. If Duff supposes that harsh treatment is likely to persuade the offender to 'sign up', then we return to the thought, supported by empirical evidence, that harsh treatment seems not to be a good strategy for achieving this.

Secondly, penitential suffering is appropriately taken up in negotiation with the wronged, regarding what burdens are required. But in state punishment there is little room for negotiation.²⁴ Further, in interpersonal relationships this negotiation leaves it open for the wrong-doer to decide to walk away. The relationship need not be restored: the parties can exit the damaged relationship. Finally, it would be deeply problematic in such interpersonal contexts for the wronged party to insist that reparative burdens are undertaken, or to coerce the wrong-doer into them.

In contrast, state punishment is not voluntarily undertaken. There is no 'walking away' when it comes to the practice of state punishment, as there is in interpersonal relationships: punitive suffering is imposed coercively. Indeed, the coercive element sometimes constitutes the suffering or burdensome element of punishment: activities that are not intrinsically painful (such as those involved in community service) are punitive *because* they are coercively imposed (see Duff, 2001, pp.99-104).

This coercive element to penitential suffering finds no support in our interpersonal practices, as Morris observes:

'the role of punishment is nonexistent, insignificant, or positively perverse in contexts where moral wrong is done a stranger or where a friendship or love relationship based on affection, respect, and trust has been damaged. [...] Infliction of punishment by the injured party upon the other party has a peculiar inappropriateness' (1971, p.430)

The role of repentance in interpersonal practice cannot support such coercive practices of punishment by the state. Our interpersonal practices involve individuals undertaking the self-imposed suffering of reparation, where this is not the only option. Punishment involves the imposition by a third party of suffering, where there is no other option.

Perhaps in political community, assumptions concerning the acceptance of appropriate burdens follow from a commitment to shared values.²⁵ As in religious communities, Duff argues, punitive suffering can be justified if members are committed to the values flouted, and if the suffering imposed is '(if not at first, then in the end) mediated by her own understanding and acceptance of its necessity' (Duff 1986, p.251). However, Duff's analogy with religious penance differs from state punishment in a number of ways.

First, membership of a state is non-voluntary and there are no exit options (especially not when one has committed a law-breaking wrong).²⁶ Membership of religious groups is not always, but sometimes, voluntary, so it may be clearer that one subscribes to the penitential rituals of that community. (Even this much is not clear: you may be a firm believer in the religion, whilst fervently seeking to reform what you regard as unduly coercive punitive measures.) Secondly, unlike being a member of a state, one also has the option of leaving a religious community: as in

24 Although, Duff notes that his communicative account favours negotiated sentencing (2001, pp.158-163).

25 Compare Murphy's claim that punitive legitimacy is secured only if citizens (could) rationally will the laws by which they are governed (1973).

26 Whilst membership of a particular state may sometimes be voluntary, membership of any state at all is not usually avoidable (and rarely beneficial).

personal relations, one can (perhaps at some cost) decide to leave rather than to take on the penitential burdens. Not so for state punishment.²⁷

Let's consider a final attempt to justify the coercive imposition of suffering. Whilst the suffering of state punishment could be embraced as a way of showing repentance, even if the wrong-doer does not sincerely repent or wish to offer a sincere apology, 'going through the motions' can nonetheless play an important role in allowing her to show that she is willing to do what is expected of her (Duff 2001, p.95, 109). Imposing punishment enables individuals to 'go through the motions' – to do what *would* be required for the repairing of relationships, if wrong-doer and wronged were to want this.

Whether or not Duff is right about the importance of ritual or formalized apologies, his proposals find no support in the processes of repentance in interpersonal relationships where, as I have emphasised, coercively imposed suffering is not permissible. To coercively impose suffering upon one's cheating partner, in order that they 'go through the motions' of undertaking penance would indeed have a 'peculiar inappropriateness'.²⁸ The ritualised aspect of apology still depends upon the willingness of the wrong-doer to undertake what is expected, if not to sincerely repent, and so cannot provide a link from the penitential suffering undertaken in interpersonal relationships to the suffering imposed by punishment.

5. Concluding remarks

The desire for suffering is a problematic component of the retributive emotions, but I have argued in support of the view that in interpersonal relationships it can be morally permissible to desire, and to aim at, the suffering of guilt. The retributive emotions, then, can be constituted by a desire for the suffering of the wrong-doer without being barbaric or base or vindictive. But this conclusion cannot provide support for retributive state punishment. State punishment imposes painful treatment and is coercive. The suffering of interpersonal relationships is of a qualitatively different sort (the pain of guilt), is induced rather than imposed, and, with respect to repentance, need not be undertaken (as one has the option of deciding not to repair the relationship). The retributivism of interpersonal relationships cannot provide justificatory support for retributive state punishment.²⁹

References

- Armstrong, K.G. 1961 The Retributivist Hits Back, *Mind* 70 (280): 471-490.
- Bennett, C. 2008 *The Apology Ritual: A Philosophical Theory of Punishment*, Cambridge University Press.
- Bennett, C. 2003 Personal and Redemptive Forgiveness. *European Journal of Philosophy* 11 (2):127–144.
- Bennett, C. 2002 The Varieties of Retributive Experience. *Philosophical Quarterly* 52 (207):145-163.

²⁷ The penitential rituals of religious communities may also differ significantly from those of interpersonal relationships, insofar as the former may not be negotiated or are coercively imposed.

²⁸ My remarks here apply primarily to adults (who may sometimes, voluntarily, go through the motions). In the case of children, some suffering (no pudding, early bedtime) can permissibly be coercively imposed in response to wrongdoing. It is unlikely that there are exit options either: the demand is not (or not usually): 'go to bed early or cease to be my child!'. But moral condemnation, blame, and penitential suffering most likely play a different role in parental-child relationships, where a significant aim is moral education and training. Reflecting on these cases shows the implausibility of extending the claim that suffering can be coercively imposed to relationships between adults.

²⁹ For detailed comments and feedback on various drafts of this paper, I am grateful to Lucy Allais, Chris Bennett, Adam Caulton, and Nathan Hanna.

- Darwall, S. 2006 *The Second-Person Standpoint: Morality, Respect, and Accountability*, Cambridge, Massachusetts: Harvard University Press.
- Duff, R.A. 2001 *Punishment, Communication, and Community*, Oxford, Oxford University Press.
- Duff, R.A. 1986 *Trials and Punishments* Cambridge, Cambridge University Press.
- Mason, M. 2003 Contempt as a Moral Attitude. *Ethics* 113 (2): 234-272.
- Brink, D. 1989 *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.
- Darwall, S. 1983 *Impartial Reason*, Ithaca, NY.: Cornell University Press.
- Feinberg, J. 1965 The Expressive Function of Punishment *Monist* 49: 397–423.
- Hampton, J. 1990 The Retributive Idea. In *Forgiveness and Mercy* Murphy and Hampton, eds. Cambridge, Cambridge University Press: 111-161.
- Hanna, N. 2008 Say What? A Critique of Expressive Retributivism. *Law and Philosophy* 27:123-150.
- Hanna, N. 2009 The Passions of Punishment *Pacific Philosophical Quarterly* 90(2): 232-250.
- Hare, R.M. 1952 *The Language of Morals*, Oxford: Clarendon Press.
- Holroyd, J. 2007 A Communicative Conception of Moral Appraisal *Ethical Theory and Moral Practice* 10 (3): 267-278
- Metz, T. 2000 Censure Theory and Intuitions about Punishment. *Law and Philosophy* 19 (4): 491-512.
- Morris, H. 1971 Guilt and Suffering. *Philosophy East and West, Symposium on Law and Morality* 21(4): 419-434.
- Murphy, J. 1990 Hatred: A Qualified Defence. In *Forgiveness and Mercy* Murphy, J. & Hampton, J. eds. Cambridge, Cambridge University Press: 88-110.
- Murphy, J. 1973 Marxism and Retribution. *Philosophy and Public Affairs* 2 (3): 217-243.
- Springer, E. 2008 Moral Feedback and Motivation: Revisiting the Undermining Effect. *Ethical Theory Moral Practice* 11(4): 407-423.
- Strawson, P.F. 1974 Freedom and Resentment. In Strawson, P.F (ed.) *Freedom and Resentment and Other essays*:.1-29.
- Tasioulas, J. 2006 Punishment and Repentance. *Philosophy* 81: 279-322.
- Tasioulas, J. 2003 Mercy. *Proceedings of the Aristotelian Society*, 103: 101-132.

