

SYMPOSIUM  
THE PHILOSOPHY OF PUNISHMENT



PUNISHMENT AND COHERENCE

BY

MICHELLE MADDEN DEMPSEY

[THIS PAGE INTENTIONALLY LEFT BLANK]

# Punishment Theory and Coherence

Michelle Madden Dempsey

Thom Brooks is to be commended for having taken on an almost impossible task in writing his impressive new book, *Punishment*.<sup>1</sup> His stated goals are ambitious: “to present a critical guide to the latest research on the leading theories of punishment and the most important alternative approaches...to consider their application in particular contexts, such as the use of capital punishment, juvenile offending, and the punishment of domestic violence, rape, and child sex offences ... and to present the most thorough explanation and defence ... to date” of his own contribution to the punishment literature, which he coins the “unified theory of punishment” (ix-x).

What is all the more impressive is that Brooks aims to accomplish these goals whilst writing in a style that is accessible to a general, non-specialist audience, thus avoiding what Americans would characterize as an “inside-baseball” approach to punishment theory. While Brooks’ efforts are mostly successful, the book does at times suffer from a lack of clarity and thoroughness.

In what follows, I will press on areas where Brooks’ *Punishment* might have benefited from further argument. This discussion is offered not so much as a critique of the book Brooks has written,

<sup>1</sup> Thom Brooks, *Punishment* (London: Routledge, 2012). Page numbers in the text are to this book.

but as an invitation to further address these underdeveloped areas in future work.

## I

### **What Counts as a Successful Theory of Punishment?**

In order to judge the success of Brooks' unified theory of punishment, or indeed any theory of punishment, we should begin with an account of what a theory of punishment is and what criteria are appropriate to evaluating its success. Brooks is more or less clear about what a theory of punishment is. Following H.L.A. Hart, he argues that a theory of punishment consists of three parts: a definition of punishment, an identification of the "general justifying aim" of punishment (or, in the case of hybrid theories, the "general justifying aims" of punishment), and an account of how punishment should be distributed (6).

With respect to his definition of punishment, Brooks again follows Hart in stipulating a definition that limits his inquiry to legally imposed punishment. The only explanation offered for this limitation is the counterintuitive view that "[i]t would be unacceptable for any individual to act in a private capacity in carrying out punishments" (5). Yet, there are many instances of private punishment that are entirely justified, such as a parent giving a "time-out" to a child who hits her younger sibling. We should not assume away the existence and justification of non-

legal punishments – nor should we presuppose that legal punishment presents the central case of punishment.<sup>2</sup>

With respect to the “general justifying aims” of punishment, Brooks is clear that his unified theory embraces a plurality of penal goals. “Punishment need not be either retributivist, deterrent, or rehabilitative, but all at once” (211). While I am sympathetic to Brooks’ pluralist approach to justifying punishment, it seems wrong to frame the issue in terms of “aims” or “goals” of punishment. Rather, we should formulate the point in terms of what punishment actually *does*, not merely what it *aims* to do. A system of punishment that aims to deter but never actually manages to deter is unjustified from the perspective of deterrence theory. A system of punishment that aims to impose deserved punishment on the guilty but only ever manages to punish the innocent is unjustified from the perspective of retributive theory. A system of punishment that aims to express condemnation and/or shape social norms but conducts its activities in secret is unjustified from the perspective of an expressive theory. And so on. I do not mean to detract from the justificatory work that can be done by having valuable aims.<sup>3</sup> My point here is simply that what punishment actually does matters as well, and that framing the issue in terms of “general justifying aims” obscures this point. We should, perhaps, instead frame the inquiry in terms of “reasons for punishment.”

<sup>2</sup> See, John Gardner, *Introduction*, HLA Hart, *Punishment and Responsibility* (OUP 2<sup>nd</sup> ed 2007), p. liii. Although, as discussed below, that legal punishment is inflicted by the state, on behalf of the community – rather than on behalf of the victim – is an important aspect of its displacement value. See n 14.

<sup>3</sup> As I’ve explained in regard to the justification of prosecutorial action, trying can have value (which I referred to as “telic value”) even when the chances of success are remote. Michelle Madden Dempsey, *Prosecuting Domestic Violence: A Philosophical Analysis* (Oxford: Oxford University Press 2009) p. 65.

Brooks has comparatively little to say on the issue of distribution as a theoretical matter, but his practical applications of the unified theory in the later third of the book offers insights to his views regarding the distribution of punishment. To his credit, Brooks resists drawing too sharp a distinction between concerns of general justification and distribution, recognizing instead that “[p]erhaps the consequences should matter” even to the question of distribution (97-98).

What remains somewhat unclear throughout the book, however, are the criteria we should use to evaluate whether any given theory of punishment is successful. We might suppose that the criteria for a successful theory of punishment simply tracks the elements of what a theory of punishment is. If so, then a successful theory of punishment will:

- (1) identify salient features of punishment (thus providing a successful definition of punishment);
- (2) illuminate considerations relevant to the justification of punishment as a general practice (thus successfully identifying its “general justifying reasons”); and
- (3) specify the conditions under which punishment may be justified in a particular instance (thereby providing a successful account of how punishment should be distributed).

Are there any further criteria in determining what constitutes a successful theory of punishment? For Brooks, the answer seems to be a resounding yes: a successful theory of punishment must be *coherent*. Indeed Brooks is concerned throughout to emphasize the coherence of his unified theory of punishment: emphasizing repeatedly that “[t]he unified theory of punishment is a unique attempt to bring together several different principles of punishment within a single and coherent approach” (123).

Unfortunately, however, it is not clear what coherence means in this context, or why it should be regarded as a necessary feature of a successful theory of punishment. If a successful theory of punishment is meant to provide an account of whether, how and why punishment is justified, then it seems sufficient to point out any and all salient features that count in favor of punishing, either as a general practice or in a particular case. We should expect these features to vary from society to society, from crime to crime, and from case to case.<sup>4</sup> In some instances, the justifications may resemble one another. In other instances, the justifications may bear little resemblance. If this is so, perhaps we should agree with John Gardner:

[Criminal punishment is] such an extraordinary abomination, that it patently needs all the justificatory help it can get. If we believe it should remain a fixture in our legal and political system, we cannot afford to dispense with or disdain any of the various things, however modest and localized, which can be said in its favour.<sup>5</sup>

Let us refer to theories of punishment that ascribe to this view as “Pick-a-Mix” theories of punishment, and distinguish them from what we might call “Coherence” theories. Pick-a-Mix theories stake no claim regarding whether any institution or particular instance of punishment is justified; they simply observe that if we wish to keep punishing, then our justification for so doing should be based on any and every consideration that

<sup>4</sup> Indeed, Brooks seems to acknowledge as much when he observes that “political societies may ... punish the same crimes in the same individual circumstances very differently in some part due to possible differences in societal contexts. Crimes and punishments may significantly differ from one political society to the next” (137).

<sup>5</sup> John Gardner, *Offences and Defences* (Oxford: Oxford University Press 2007), p. 214.

weighs in its favor.<sup>6</sup> If something can be said in favor of punishment, then (according to the Pick-a-Mix theories) we should include that consideration in our justification of punishment. This is true even of considerations that weigh only very weakly in favor of punishment, and of considerations that weigh in favor of only some kinds of punishments but not others. It is also true of considerations that do not manifest any degree of theoretical coherence. For the Pick-a-Mix theorist, worrying about the degree of coherence amongst considerations that bear on the justification of punishment is simply a waste of intellectual energy. We should instead dedicate ourselves either to identifying salient features of punishment that genuinely weigh in its favor, or set ourselves the task of abolishing punishment for lack of justification.

Brooks rejects this approach to punishment theory, illustrating his disdain through a detailed criticism on the Model Penal Code sentencing principles, which simply list multiple goals of punishment without attending to concerns of how these various considerations cohere.<sup>7</sup> Brooks' complaint against Pick-a-Mix theories of punishment (as illustrated in the Model Penal Code) is not that it draws upon multiple penal goals. Indeed, his unified

<sup>6</sup> Gardner, n 5, p. 214.

<sup>7</sup> The Model Penal Code approach is embodied in §1.02:

(2) The general purposes of the provisions governing the sentencing and treatment of offenders are:

- a. to prevent the commission of offences;
- b. to promote the correction and rehabilitation of offenders;
- c. to safeguard offenders against excessive, disproportionate or arbitrary punishment;
- d. to give fair warning of the nature of the sentences that may be imposed on convictions of an offence;
- e. to differentiate offenders with a view to a just individualization in their treatment.



theory of punishment is similarly pluralistic. The problem, Brooks argues, is that simply placing “multiple penal goals [in] a list lack[s] a suitably robust framework that offers a sufficiently clear steer on how these goals relate to one another within the framework” (132). He explains:

Why should any of these goals be included? The answer seems to be that each is intuitively attractive on its individual merits. But this fails to address specifically how each might relate. Imagine making a cake combining only those ingredients that you enjoy individually. Following this procedure may not guarantee that all the necessary ingredients for making a cake are included. Nor is there any guarantee that the cake will be edible. Now imagine starting a company by inviting only those persons that you enjoy working with individually. This procedure may not guarantee that all the necessary tasks will be covered. Nor is there any guarantee that the company’s members will work together suitably effectively. These examples centre on the problem of justifying a legal practice without sufficient consideration of how the individual parts coherently work together in support of the practice aims (132-133).

But what is the special ingredient in Brooks’ theory that makes the plurality of penal goals he endorses cohere any better than the laundry list of reasons on offer in a Pick-a-Mix theory? What is it that makes the unified theory unified? Brooks’ explanation is opaque. He claims that “[t]he unified theory of punishment overcomes this problem” of incoherence because “[i]t addresses desert, proportionality, and other penal goals [as] they come together within a larger unified framework” (133). To this point in his explanation, we must take it on trust. The unified theory is unified because Brooks keeps telling us it is.

Yet, how does this unity manifest itself? How does the coherence of the unified theory inform the way we think about punishment, so that our thinking is different, better than it would be under a Pick-a-Mix theory? Brooks offers the following response: “[Under a unified theory approach] we don’t weigh up

possible sentences in light of general deterrence versus desert and other penal considerations because we find them intuitively attractive individually” (133). I confess to not understanding what Brooks means at this point. Does he mean that we don’t weigh up possible sentences in light of general deterrence versus desert and other penal considerations at all – or merely that we don’t do so because we find them intuitively attractive individually?

The remainder of Brooks’ explanation does little to clarify the point: “Punishment does not bring together multiple penal goals because it can, but because it should. Punishment is a response to crime that aims at the restoration of rights. Punishment addresses multiple penal goals in serving its aims” (133). While I have no disagreement with any of these claims, it remains unclear how they explain what it means to say that the unified theory is coherent in a way that makes its penal pluralism more attractive than other hybrid theories of punishment.

Any plausible explanation would have to point to one or more penal goals (reasons for punishment) that play a cohering role in the theory’s account of how and when punishment is justified, and explain how each goal hangs together in a coherent whole. Brooks offers no account of how this cohering relation between multiple penal goals is achieved under the unified theory. Yet, there is such an explanation available to another hybrid theory that Brooks rejects: expressivism.

## II

### **The Coherence of Expressivist Theories of Punishment**

An expressivist theory of punishment can provide the coherent penal pluralism Brooks prizes in the following way.

First, the theory will identify multiple operative reasons for punishment similar to the list that Brooks endorses: retribution, deterrence, etc.<sup>8</sup> All are operative reasons for punishment insofar as they are values that can be realized through punishment - each operative reason is capable of doing some normative work in justifying punishment. Second, the expressive theory will identify how the expressive function of punishment provides auxiliary reasons for punishment that relate to each of operative reasons.<sup>9</sup> The expressive function of punishment provides auxiliary reasons because it identifies punishment as an act which there is reason to perform under the circumstances. Which is to say, even if the expressive function of punishment is neither a complete reason, nor even an operative reason for punishment, the expressive function of punishment nonetheless helps to identify punishment as a justified response to the defendant's crime for reasons of deterrence, retribution, etc.—and as such the expressive function of punishment serves as an auxiliary reason for punishment.<sup>10</sup>

<sup>8</sup> On operative reasons, see Joseph Raz, *Practical Reason and Norms* (Oxford: Oxford University Press 1990), pp. 33-34.

<sup>9</sup> On auxiliary reasons, see Raz, *Practical Reason and Norms*, pp. 34-35. Raz describes two roles auxiliary reasons play in practical reasoning: identifying and strengthening. Identifying auxiliary reasons “help identify the act which there is reason to perform.”

<sup>10</sup> I take no view here as to whether the expressive function of punishment is also an operative reason – that is, whether there is a value in expressing whatever it is that punishment expresses. In previous work, I have argued that such value exists and grounds operative reasons for prosecutors to pursue (or not pursue) certain kinds of prosecutions. My point here is that even if the expressive function of punishment does not have value, it does help identify punishment as the means for realizing other values such as deterrence, retribution, etc. – that is, it functions as an auxiliary reason of the identifying type. Raz, *Practical Reason and Norms*, pp. 34-35. The expressive function of punishment may also serve as an auxiliary reason of the strength-affecting type. I assume as much in the discussion in the main text.

As the expressivist theory we are examining is pluralistic, it can admit of multiple operative reasons for punishment – deterrence-based operative reasons, retributive-based operative reasons, displacement-based operative reasons, etc.<sup>11</sup> The relationship between these various operative reasons can be illuminated by explaining how each relates to the expressive function of punishment: specifically, the fact that punishment is expressed/communicated as the intentional infliction of a loss *for breaking the law* (that is, its expressive function) is an auxiliary reason that picks out punishment as a particularly effective way to realize deterrent, retributive, and displacement value. I will consider each relationship in turn below.

The expressive function of punishment serves as an auxiliary reason relating to deterrence-based reasons in favor of punishment, because it helps to identify punishment as a particularly effective way to realize the value of deterrence. As many have argued and as Brooks agrees, deterrence is indeed an important value that can be realized by punishment – which is to say, deterrence is an operative reason in favor of punishment. Yet, the deterrent effect of punishment depends to a significant degree on the fact that the punishment is *expressed* to the defendant and potential future defendants as the intentional infliction of a loss for breaking the law. Imagine a punishment that is not communicated to the defendant. The state imposes a fine in response to the defendant's crime, but it does so not by public declaration of the punishment, but instead by simply withdrawing the funds from the defendant's bank account or sneaking into his home to obtain the cash. Absent the public expression of the punishment, the defendant is likely to be confused, not deterred. Similarly, such a punishment would be

<sup>11</sup> On displacement, see Gardner, *Offences and Defences*, pp. 213-216, and the text accompanying nn 13-16 below.

incapable of achieving general deterrence. Unless the punishment is communicated to potential future defendants as the intentional infliction of a loss for breaking the law, they are far less likely to be deterred by punishment. In this way, the expressive function of punishment serves as the handmaid of a deterrence-based justification of punishment.

Now consider retribution-based operative reasons in favor of punishment and how they relate to the expressive function of punishment. Retributive theories are correct in supposing that there is value in a defendant suffering some intentionally inflicted loss for having committed a crime.<sup>12</sup> Which is to say, retributive value is something that can be said in favor of punishment. Yet, if punishment lacks its expressive function, then the defendant will not register the value of suffering *for his crime*. If the state imposes a secret punishment, making the defendant's life go less well in any variety of ways, but never communicates to the defendant that this treatment is being imposed *for his crime*, then the retributive value of the punishment not realized. Just as with deterrence, the expressive function of punishment serves as the handmaid of retributive justifications of punishment.

There is a similar story to be told with respect to displacement-based operative reasons in favor of punishment. As John Gardner puts it:

That people are inclined to retaliate against those who wrong them, often with good excuse but rarely with adequate justification, creates a rational

<sup>12</sup> I would rather put the point in terms of “for having committed a wrong that also happens to be a crime” to screen out cases in which a defendant is punished for having committed a crime that is not also a wrong. Brooks frames the point in terms of crime, so I will follow suit here.

pressure for social practices which tend to take the heat out of the situation and remove some of the temptation to retaliate.<sup>13</sup>

This “rational pressure” is what I have referred to as displacement-based operative reasons in favor of punishment. Its displacement value is indeed something that can be said in favor of punishment. Yet, again, the expressive function of punishment is key to realizing punishment’s displacement value. Here, we should expand our understanding of the expressive function of punishment beyond the characterization offered above, and layer in the fact that *legal punishment* is imposed by the *state* – acting on behalf of the community – rather than being imposed by victims or on behalf of victims.<sup>14</sup> By imposing legal punishment, the state (community) expresses to potentially vengeful victims something along the lines of, “Just chill ... we’ve got this. We will address this crime adequately – so that you will have no reason to take matters into your own hands.”<sup>15</sup> To the extent that the state fails to express that *it alone* is the primary agent in inflicting punishment, the displacement value of punishment is reduced. Moreover, to the extent that the state fails to make good on its promise of an adequate response to crime, it fails to displace reasons victims may have to take matters into their own hands. Unfortunately, the criminal justice system often stumbles on both

<sup>13</sup> Gardner, *Offences and Defences*, p. 214.

<sup>14</sup> Above (n 2), I noted that Brooks was wrong to suppose that legal punishment was the central case of punishment. Still, if a key pillar of the justification of legal punishment lies in its displacement value, then the fact that the state imposes punishment is an important feature of the justification of *legal* punishment.

<sup>15</sup> Some criminal law abolitionists have argued that the state stepping in to “steal” victim’s conflicts in this way counts against legal punishment. Nils Christie, *Conflicts as Property*, 17 *British Journal Of Criminology* 1-15 (1977). Yet, to the extent that the expressive function of punishment helps to realize a displacement value of punishment, the abolitionist argument is weakened.

fronts: it too often turns its discretionary authority over to victims and fails to provide an adequate response to crime.<sup>16</sup> Still, if a system of punishment is functioning properly, the expressive function of punishment serves as the handmaid of displacement justifications of punishment.

A similar explanation may be offered with respect to other values realized by punishment, but I hope to have done enough to motivate the possibility that an expressivist theory of punishment is capable of providing the coherent penal pluralism Brooks prizes. In sum, such a theory not only explains how punishment can be justified in terms of a plurality of operative reasons in favor of punishment, it illustrates how the expressive function of punishment serves an auxiliary reason that identifies punishment as a particularly effective means of realizing the values that ground those operative reasons.

Moreover, an expressivist theory is attractive insofar as it informs our practices of punishment by highlighting the importance of transparency and publicity in our punitive practices. If we sacrifice transparency and publicity (that is, if we compromise the expressive function of punishment), it becomes all the more difficult to justify our punitive practices. Interestingly, this point holds true even when our punitive practices are beyond reproach – even when we are only punishing the deserving, and only for serious wrongs that cause substantial harm. Often we think of transparency and publicity as important only for uncovering official corruption or discovering and checking the misguided exercise of official discretion. Yet, as the expressivist theory I’ve outlined above demonstrates,

<sup>16</sup> These failures are most starkly illustrated in domestic violence prosecutions, where victims’ stated wishes are often treated as authoritative, while the state nonetheless fails to provide an adequate response to these crimes. See, Dempsey, *Prosecuting Domestic Violence*, ch. 9.

transparency and publicity are central to the justification of punishment even for an otherwise perfectly well-functioning system.

### III

#### Conclusion

In conclusion, Thom Brooks' *Punishment* provides an intriguing and insightful account of punishment and its justification. I find his theory of punishment particularly appealing in virtue of its embrace of penal pluralism. My reflections have focused upon Brooks' search for coherence amongst the plurality of reasons that may weigh in favor of punishment. If the concerns I've raised are correct, this search for coherence may be unnecessary (that is, perhaps the Pick-a-Mix theorists are correct to think that searching for such coherence is a waste of rational energies). If, however, coherence *is* an important aspect of a pluralistic justification of punishment, then Brooks may do well to reconsider whether the expressive function of punishment can provide that coherence, by unifying otherwise disparate, pluralistic reasons for punishment through its role as an auxiliary reason.

*Villanova University*



If you need to cite this article, please use the following format:

Dempsey, Michelle Madden, "Punishment and Coherence," *Philosophy and Public Issues (New Series)*, Vol. 5, No. 1 (2015), 43-56, edited by S. Maffettone, G. Pellegrino and M. Bocchiola