Philosophia (2016) 44:1099–1124
DOI 10.1007/s11406-016-9773-0

Punishment, Forgiveness and Reconciliation

Bill Wringe[1]

Received: 22 July 2016 / Revised: 26 September 2016 / Accepted: 29 September 2016 /
Published online: 12 November 2016

Abstract It is sometimes thought that the normative justification for responding to large-scale violations of human rights via the judicial appararatus of trial and punishment is undermined by the desirability of reconciliation between conflicting parties as part of the process of conflict resolution. I take there to be philosophical, as well as practical and psychological issues involved here: on some conceptions of punishment and reconciliation, the attitudes that they involve conflict with one another on rational grounds. But I shall argue that there is a conception of political reconciliation available which does not involve forgiveness and this forms of reconciliation may be the best we can hope for in many conflicts. Reconciliation is nevertheless likely to require the expression of what Darrell Moellendorf has called 'political regret' and the denunciatory role aspect of punishment makes it particularly well-suited to this role.

Keywords

1 Introduction

My aim in this paper is to say something about the nature of forgiveness and reconciliation in the public domain and their relation to the practice of criminal punishment. I shall be particularly concerned with questions about institutional responses to large-scale public wrongdoing, and with the questions of whether and to what extent the projects of punishing wrongdoing and of achieving reconciliation between different parts of society are in tension with one another. By 'public wrongdoing' I mean wrongdoing which is done on behalf of, or purportedly on behalf of the public, via the organs which are supposed to act and speak on their behalf. In other

✉ Bill Wringe
wringe@bilkent.edu.tr

[1] Department of Philosophy, Bilkent University, Ankara, Turkey

words I shall be concerned by wrongs done by the state and by organizations, both formal and informal which can be taken to be acting on behalf of the state.[1] I shall also be concerned primarily with the case of legal punishment performed by state institutions, rather than supra-national or international organizations, or such bodies as the International Criminal Court.[2]

The kind of wrong-doing that I am interested in is large-scale public wrongdoing. Largeness of scale here is intended to refer both to the extent of the wrongdoing involved and its seriousness. So I intend to ignore both wrongs committed against isolated individuals (except insofar as these form part of a more general pattern) and more pervasive small-scale wrongs (such as, say a general pattern of petty thievery on the part of government officials.) I shall also be concerned with crimes committed against individuals who are – or ought to be regarded as – members of the society which is being discussed. So my focus will include – as paradigmatic examples – crimes committed under apartheid by the Republic of South Africa against black South Africans; crimes committed by the British state in Northern Ireland; many instances of genocidal killing in Rwanda, and, no doubt, many others.[3]

The concept of reconciliation is a complex one: different authors have given accounts of what reconciliation requires which seem incompatible with one another. We might even suspect that different notions of reconciliation are appropriate in different contexts: for example political reconciliation might turn out to be significantly different from interpersonal reconciliation.[4] I shall examine the concept of reconciliation – and in particular its relationship with the notion of forgiveness in more detail later in the paper. Nevertheless, it may be worth mentioning some constraints on what can count as an appropriately political conception of reconciliation. I take there to be at least three such constraints. First, it must be a state which is in principle achievable, and achievable by recognizably political (rather than, say, religious or spiritual) means. Secondly, it must have some kind of public dimension : even if it does not involve acts performed on the part of the public, it must involve acts performed in ways that are open to the view of the public or states of affairs which are similarly open to it. Thirdly, it should not presuppose any kind of religious or spiritual outlook which cannot be expected to be widely shared across different groups in society.

It is easy to see that the goals of punishment and reconciliation are distinct; that what promotes one may not promote, and may even stand in the way of the other; and that resolving such conflicts will often require imaginative and creative

[1] Wrongdoing can be public in a number of different senses: it can be committed in public (rather than behind closed doors); by individual members of the public; or be properly of concern to the public. (Duff 2001 has argued that publicity in this last sense is definitive of behavior which deserves criminalization) Much of the wrongdoing with which I am concerned with here will be public in at least some of these other senses.

[2] I shall make the reasons for this limitation explicit in Section X: they have to do with the extent to which international organisations can be regarded as speaking for the public. For a discussion of how the denunciatory theory of punishment which I defend here might apply to cases such as punishment by international tribunals and by the ICC, see Wringe 2006, 2010 and Wringe 2016a Chapters 7 and 8.

[3] If, as Vernon 2002 argues, the use of state power against citizens is definitive of the category of crimes against humanity, many though perhaps not all of the examples that I shall be concerned with fall into that category.

[4] As a number of authors, including Moellendorf 2007 and Murphy 2010, whose work I draw on below suggest.

political leadership.[5] It is less obvious why we might expect philosophers to be able to say something illuminating about these issues.[6] However, I shall argue first that the projects of punishing wrongdoing and achieving reconciliation are in tension with one another for reasons that derive from apparently fundamental features of the ways in which we think about both punishment and reconciliation; and subsequently that these tensions can be dissolved by more careful consideration of what is constitutively involved in both punishment and political reconciliation.

A brief overview may be helpful. Sections II-V are concerned with the apparent tension between punishment and reconciliation. In Section II, I give an argument for taking punishment to have an expressive dimension. In Section III, I discuss one especially well-known view of what punuishment might be taken to express: Feinberg's idea that punishment expresses vindictvie resentment. In Section IV I consider reconciliation: I argue that it involves a form of acceptance of wrongdoing, and that it in the cases we are interested in it should not be seen as a state which only involves attitudes of citizens to one another. In Section V I draw these strands together: I argue that punishment and reconciliation involve opposed attitudes and that these opposed attitudes must be attributed to the same subject, namely the state.

Sections VI-X consider ways in which an expressivist might resolve the apparent tension for which I have argued. In Section VI, I outline a response based on Antony Duff's 'communicative' version of expressivism. In Section VII, I argue that this response fails, because it overstates the role which forgiveness plays in reconciliation. In Sections VIII I argue that a better conception of political reconciliation is available, and that on this view we should see the expression of political regret as partly constitutive of political reconciliation. This allows for a different way of resolving the tension between puinshment and reconciliation, based on a different kind of expressivist theory of punishment. I explain and argue for this version of expressivism, which I call 'denunciatory expressivism', in Section IX.[7] Finally, in Section X I explain how, on the denunciatory view, punishment can express political regret for state-sponsored wrongdoing, and thus how it can be partly consititutive of political reconciliation.

2 Punishment as Expressive

In recent years, many authors have argued that punishment has an important expressive dimension.[8] Some expressivists have argued the practice of punishment is best

[5] As both Minow 1998 and May 2004 emphasize. In a discussion of debates about the resolution of conflicts in the former Yugoslavia in the 1990's, Minow notes a concern among policy-makers at the time that 'to proceed with arrests and trials of leaders would undermine efforts not only to negotiate, but to stabilize efforts for justice and peace (Minow 1998 p 39) Similarly, May writes 'The goals of reconciliation all concern the attainment of that peace and security among people that we all seek…criminal trials sometimes exacerbate rather than diminish the tensions and divisions between peoples.' (May 2004 pp 237–8)

[6] As one anonymous referee for this journal made especially clear.

[7] I there note in particular how some of the considerations which Feinberg takes to support his version of expressivism seem to support my view as well, and argue that for those who are committed to some form of expressivism on the basis of the arguments put forward in Section II should accept the denunciatory view rather than any other currently articulated version of expressivism.

[8] See for example Feinberg 1965; Duff 1986, 2001; Falls 1987; Primoratz 1989; Hampton 1992; Metz 2000, 2007; Bennett 2008; Glasgow 2015 UTHOR (various) Wringe 2006, 2010, 2013, 2015, 2016a

explained by the fact that the state is under an obligation to express certain kinds of values to its citizens, and that punitive sanctions provide one – and perhaps the only appropriate means of expressing those values.[9] We might also think that communicative – and some denunciatory – forms of expressivism can explain how punishment can avoid treating punished offenders as mere means; that some forms of expressivism can make sense of the intuitions that support retributivism[10]; and that expressivism can explain features of our penal practice such as the high burden of proof in criminal trials.[11]

Authors who have emphasized punishment's expressive dimension have differed on a number of further issues. Three such issues are particularly worth noting. The first is the question of whether the expressive dimension is constitutive of a form of treatment's being punishment, or merely a feature which punishment typically possesses. The second is what, exactly, we should take punishment to express. The third is whether the expressive dimension plays an important role in explaining how punishment can be justified. In this paper, I shall mostly be concerned with the first and second of these questions, rather than the third. In this section I shall address the first question : I shall argue that we have good reason to take the expressive dimension of certain forms of hard treatment to be constitutive of their being punishment. By 'we' I mean, in particular, anyone who regards the question of whether punishment can be justified to merit serious enquiry. The question of what, in particular, we might take punishment to express will be addressed in Section III, and a number of subsequent sections.

I shall start by arguing for a view which I shall call Unspecific Constitutive Expressivism or UCE for short.

> UCE: It is a necessary condition of a way of treating convicted offenders constituting punishment that it have some kind of expressive dimension, (which may be further specified by any given theorist.)

UCE is unspecific in the following sense: it does not make any commitment to the kind of expression which punishment must involve. It is therefore compatible with a wide range of expressivist views, including those of Feinberg, Duff, Bennett, Metz and myself. (It should be distinguished from a view with which it might easily be confused, but which many of these advocates of expressivism would deny - and which is of dubious coherence: namely, that the content expressed by punishment must be something which is unspecific.) Furthermore I shall argue for UCE in a way that remains neutral between various views as to what exactly punishment might express.

UCE by itself says nothing about the ways in which the institution of punishment might be justified. Although many of those who are committed to UCE might take the expressive dimension of punishment to play a role in justifying punishment, one might also accept UCE, while denying that punishment is justified at all. This is worth noting here because many arguments which are labelled 'objections to expressive theories of punishment' are in fact objections to claims about the ways in which the expressive dimensions of punishment can justify the institution of punishment. Consider, for

[9] Metz 2000, 2007; Glasgow 2015
[10] Duff 2001; Wringe 2006
[11] Duff 2007

example, the so-called 'problem of hard treatment'. This problem typically takes the following form: given an expressive feature of punishment which is supposed to play a role in explaining why punishment is justified, the proponent of the problem asks why it would not be possible to replace punishment with an alternative response to crime which had a similar expressive function, but which did not involve inflicting hard treatment on offenders. This is a significant problem for those who think that expressive considerations play a role in justifying punishment. But it need not be a problem for advocates of UCE since they need not be interested in claiming that the expressive features of punishment explain why punishment is justified.[12]

This is not to say that UCE, or the argument I shall give for it, is entirely irrelevant to questions about justification. Even though UCE says nothing about how the expressive features of punishment explain why punishment is justified, it does show us how to resist an argument which purports to show that punishment cannot be justified. In fact, I shall argue that those of us who think that punishment is sometimes justifiable should accept UCE precisely because it enables us to resist this argument.[13]

The argument I have in mind has been articulated most clearly by David Boonin. It starts from the premiss that punishment must involve an intention to harm offenders and claims that none of the supposed justifications of punishment explain how this can be permissible.[14] As I have argued elsewhere, someone who is committed to UCE can avoid commitment to the key premiss, in a way that those who do not accept UCE cannot.[15] For expressivists can claim that in order to be punishment a form of treatment must involve harsh treatment inflicted in response to wrongdoing, by a legitimate authority, understanding harshness in the following way: treatment is 'harsh' if and only if it is treatment of a sort which can be reliably expected to cause suffering to individuals of the same type as the individual on whom it is inflicted. On this way of seeing things, one can inflict harsh treatment – and in particular forms of harsh treatment which constitute punishment - without intending to cause suffering.

Why is this move available to advocates of UCE, but not to advocates of accounts of punishment which do not incorporate an expressive element? The answer is that on this account of harshness pre-trial detention, and legally sanctioned self-defense will all be forms of legally sanctioned harsh treatment in response to an offense. An expressivist can deny that these legally sanctioned forms of harsh treatment are punishment, on the grounds that they lack the expressive elements which they take to be essential to punishment. But those who do not accept UE cannot make this move. They will

[12] Ref Glasgow 2015; Wringe 2013

[13] Someone might think that in appealing to UCE to resist Boonin's argument for the view that punishment can't be justified, I acquire an obligation to respond to the 'hard treatment problem, since I am making the expressive nature of punishment relevant to the justification of punishment. But I do not think so. Appealing to UCE in order to show that a premiss in Boonin's argument is false is not the same as appealing to some particular expressive feature of punishment explains why punishment is justified. We can see this by noting that someone who accepted UCE might think it undermines Boonin's argument, while thinking that any expressive feature of punishment that might conceivably explain why punishment is justified must fail to do so because that feature of punishment could be shared by other responses to crime that did not involve harsh treatment. One could even think all of this while holding that the harsh treatment involved in punishment might be justified is some way entirely independent of it's expressive features (for example, on deterrent grounds)

[14] Boonin 2008;

[15] Wringe 2013

thereby be committed to the view that forms of treatment which we would normally regard as significantly distinct from punishment are forms of punishment.

This argument need not compel one to accept UCE. It is worth noting two reasons for this, and responding to them. First someone might appeal to the possibility of finding some way of distinguishing punishment from harsh treatment inflicted by a legitimate authority in response to an offense that drew on some feature or features of punishment other than its expressive dimension. Nevertheless, it seems reasonable to expect someone who wishes to rely on it to avoid commitment to UCE to give some suggestion of what aspect of punishment, other than its putatively expressive nature, they would appeal to in order to show that pre-trial detention and legitimate self-defense were not forms of punishment. However, In the absence of specific suggestions as to what these features might be it seems reasonable to set this objection aside. In doing so, we give up the ambition to provide an argument which establishes its conclusion with irrefutable certainty. But this is an ambition which philosophers rarely realise in any area; and it seems especially inappropriate here.

Secondly we might note that the argument appears to assume what we might, in other contexts be expected to regard as controversial: namely, that punishment can indeed be justified. I set it aside in this context for two reasons. First, the argument for UCE does not, in fact, rest on the claim that punishment can be justified, but simply on the claim that we should not take the Boonin-style argument to show that it is not. (One thing which should make us suspicious of that argument is that it shows not only that punishment cannot be justified but that it is *obvious* that it cannot be justified.) Secondly, although there are contexts where it make sense to raise questions as to whether the institution of punishment can be justified at all, there are also contexts in which it is justifiable to put such concerns aside. Situations where we are considering state responses to large-scale wrongdoing seem to provide one such context. Those who think that punishment can never be justified are not likely to be serious participants in debates of this sort.

3 Punishment as Expressive of Vindictive Resentment

In a famous essay, Joel Feinberg has claimed that punishment involves the expression both of 'judgments of disapproval' and of 'a kind of vindictive resentment'; or as he later puts it 'a kind of fusing of resentment and reprobation'.[16] We should notice that Feinberg does not simply claim that these attitudes are frequently associated with punishment, or that those who inflict punishment on others often experience these attitudes. Instead, his view is that where these attitudes are not present we have something distinct from punishment in what he calls 'the strict and narrow sense'.[17]

Feinberg's view goes beyond UCE, insofar as it makes a specific claim about what punishment expresses: namely, a fusing of vindictive resentment and judgments of reprobation. Call what Feinberg adds to UCE 'Vindictive Resentment Expressivism' or VRE for short. VRE embodies one particular view of what punishment expresses, which one might find plausible, but which some expressivists might deny. As we shall

[16] Feinberg 1965 pp 402–3.
[17] Feinberg 1965 pp 397–8.

see, it is VRE, rather than UCE, which generates the apparent tension between punishment and reconciliation with which I shall be concerned. Eventually, we will have to reject it, and replace it with a more adequate view as to what punishment expresses. But it is worth considering what might be said in favor of VRE before we do so.

Feinberg starts by noting 'that the expression of the community's condemnation is an essential ingredient in legal punishment is widely acknowledged by legal writers'. He cites the American writer Henry Hart (who in turn cites a Professor Gardner in support of his view) in support of this proposition. He cites, but does not comment on Hart's suggestion that we might appeal to the idea that certain penalties express societal condemnation in order to pick out what is distinctive about criminal – as opposed to civil – law; and also appeals to what he takes to be phenomenological features of the prison experience, saying that 'to any reader who has in fact spent time in a prison…'hatred, fear or contempt for the convict will not seem too strong an account of *what punishment is universally taken to express*' (my italics)

Feinberg thus takes the idea that punishment involves some kind of disapproval to be something of a common-place, both among legal writers and the wider public. This does not seem enough to establish that such judgments are definitive of punishment. However, if we already have reason to think that expressive elements play a role in the definition of punishment (as the arguments of Section II suggest), we might take Feinberg's comments to provide us with an illuminating account of what it is that punishment expresses. But Feinberg also goes further, and suggests that it is the fact that punishment has this kind of expressive role which further explain what he calls 'derivative symbolic functions' of punishment. These functions include the authoritative disavowal of certain forms of behaviour; symbolic non-acquiescence in that behaviour; vindication of the law; and absolution of non-criminals.

Feinberg says of these functions of punishment they 'presuppose the expressive function and would be impossible without it'. Since the functions Feinberg has in mind are themsleves expressive, it seems hard to deny that they presuppose that punishment has some sort of expressive function. We might nevertheless doubt whether they require 'vindictive resentment' directed at the criminal, rather than say merely authoritative denunciation of an offender's acts. However, what we do seem to be able to say is that if we have independent reason for thinking that punishment expresses vindictive resentment - as Feinberg's appeal to the experience of those who have spent time in prison suggests – then the fact that punishment has this kind of expressive nature will allow it to perform these expressive functions. This will not show that punishment must have this expressive function. It will, however, suggest that if we think the derivative symbolic functions of punishment are worth preserving there is at least some reason for allowing punishment to express vindictive resentment.

4 What Reconciliation Involves

So much, at least for now, for the attitudes associated with punishment. What should we say about the attitudes involved in reconciliation?

It is natural to think of reconciliation between the victims of wrongdoing and its perpetrators as involving the acceptance of the perpetrators of wrong-doing, made in full knowledge of the wrong that they have done, by the victims of wrongdoing. The

term 'acceptance' functions as something of a place-holder here: on different accounts of reconciliation, it might be filled out in different ways. On some conceptions of reconciliation, it might involve forgiveness of the wrongdoer. But on other conceptions of reconciliation, it may involve less : for example, a willingness to regard those found guilty of wrongdoing as sharing in a common political life.[18] There is, nevertheless, some kind of 'lower bound' on what we might plausibly take 'acceptance' to involve here: merely acknowledging the continued existence of wrongdoers and allowing their existence to continue unmolested on prudential grounds would not be a form of acceptance which would be even partially constitutive of reconciliation on any plausible view of what reconciliation might involve.

Can the kind of reconciliation we are interested in achieving in states which have been marked by large-scale wrongdoing be thought of as something which only involves relationships between individual citizens? I suspect not. There are two reasons for doubting it. First, it may be unrealistic, in states with a recent history of large-scale public wrong-doing, to think that individuals can be reconciled with one another without the public expression of support for the attitudes which are constitutive of reconciliation. If so, then even someone who thinks that the kind of reconciliation we are looking for in states which have been marked by large scale wrongdoing involves widespread interpersonal reconciliation may have a reason for supporting an account of reconciliation which has a non-political dimension.

But there is in any case at least some reason for thinking that the kind of reconciliation which we ought to aim for in states which have been marked by large-scale public wrongdoing requires something less than widespread interpersonal reconciliation. Achieving widespread interpersonal reconciliation in the wake of large-scale wrongdoing is a goal which may be implausibly demanding, morally problematic, or politically inappropriate. It may be impossible for every individual victim of wrongdoing in such a society to become reconciled with every individual wrongdoer. Such a state of affairs may require more than we can reasonably demand of those who have been wronged. It may also simply require wronged individuals to entertaining reconciliatory attitudes to an implausibly large number of individuals. Even if it is not impossible for every wronged individual to be reconciled with every wrongdoer, it may not be morally appropriate. Perhaps there are some wrongs which those who suffer them should not be willing to accept, even when they are in a position to be certain that those wrongs will not be repeated. And finally it may be wrong to treat the achievement of a state of affairs in which each individual is reconciled to each other individual as a political goal: it might, for example, be inappropriate for a liberal state to seek to 'get inside the head' of the victims of wrongdoing in the ways that would be required if we took this as a political goal.

A political conception of reconciliation – of the sort which I shall explore in Sections VI-VII – need not be objectionable in this way. For it may – and the conceptions of

[18] Forgiveness is itself a somewhat contested term. It is often characterized (following Butler) as the giving up of resentment. If we understand the term resentment as a place-holder for a variety of vindictive attitudes, as Murphy 2003 suggests, then we might think that even on this view, we should see reconciliation as involving forgiveness. If so, and if we accept the Feinbergian view that punishment involves a kind of vindictive resentment then the tension between the attitudes involved in punishment and those involved in reconciliation will seem especially clear. However I shall argue in Sections VII and VIII that it is helpful to think of reconciliation in somewhat different terms.

political reconciliation which I look at in those sections in detail do - demand less of individual citizens than that each of them become reconciled to each of their fellows. But if it is not individual citizens which are the subject of the attitude of reconciliation on a political conception of reconciliation, then something else will have to be. Even if the state is not the inevitable substitute here, it is perhaps the most likely: the acceptance of wrongdoing may be signaled by public statements, by acts of amnesty (though if the arguments of this paper are correct, these will be problematic), or in many other ways.[19]

5 The Tension Between Punishment and Reconciliation

In Section IV I argued that the reconciliation involves some kind of attitude of acceptance. However we understand the notion of acceptance in this context, it seems to be one which is fundamentally at odds with the attitude which Feinberg took to be characteristic of punishment – namely vindictive resentment. We might also note that the conflict between these attitudes is not simply a matter of psychology: a question of human beings finding it difficult to hold certain one set of attitudes in conjunction with another one. Instead there seems to be a rational conflict between the two sorts of attitudes. Feeling vindictive resentment against someone seems to be rationally precluded by whatever kind of acceptance is partly constitutive of reconciliation.[20]

This being so, we might initially suppose that if Feinberg was right in taking the view that punishment expresses vindictive resentment to be widely - if perhaps somewhat inchoately – shared, this would be enough to explain a fairly widespread sense that there is a tension between punishment and reconciliation. We might also think that if Feinberg was right to think that this view of the expressive nature of punishment was not merely widely shared, but actually correct, then we would be correct in taking there to be a tension between punishment and reconciliation.

However, matters are not quite so simple. For someone might suppose that the attitudes involved in punishment and those involved in reconciliation could be attributed to different subjects. If so, their coexistence would give rise to no more philosophical discomfort than the coexistence of my belief that Bach was a sublime musical genius with my sister's profound distaste for his whole oeuvre. But, as I shall now argue, this tempting suggestion fails. In the kinds of case we are interested in, we cannot attribute the attitudes involved in punishment and those involved reconciliation to distinct subjects.

[19] An alternative possibility is that we should see reconciliation as a relationship between different groups in a society. However I am skeptical about this suggestion. The proposal only seems promising if we think of the kinds of groups to which it applies being groups to which membership is assigned on the basis of identity-based features such as ethnicity, race or religion. However, although I am committed to the view that some kinds of groups - for example, agent-like groups such as states, and Gilbert style plural subjects - can be the subjects of the kinds of attitudes that I take to be required for reconciliation, this will not in general be true of identity-based groups. For further discussion, see Section VI below.

[20] People can sometimes hold attitudes that conflict in the ways I take the attitudes involved in punishment and reconciliation to conflict. But people can also have inconsistent beliefs: this doesn't show that there is no rational conflict between believing p and believing not p. For helpful discussion of the phenomenon in the emotional case, which emphasizes that there is indeed some kind of conflict between the content of the attitudes involved see Greenspan 1980.

Let us start by considering what an advocate of UCE should say about the attitudes involved in punishment. UCE does not entail any particular view about what punishment expresses. Still one might think that if punishment expresses anything, one thing it is likely to express is some form of disapproval of a criminals actions.[21] One might then take UCE to be refuted by the fact that an individual might commit a crime and be tried, sentenced and serve their sentence without any individual involved in the process intending to express disapproval of the sentenced individual's actions. The judge might feel the form of behavior dealt with by the law in question should be decriminalized; the jurors may be indifferent; and those involved in jailing the prisoner might regard their role in the penal system as simply one job among many.

However, these points do not constitute a refutation of UCE. They do suggests that advocates of UCE should be more careful than many expressivists have been about who exactly they take to be expressing what punishment expresses. Where legal punishment is concerned, they should take the subject of those attitudes to be the state, conceived of as a collective agent.[22] Although the attitudes of the expressed by the state will depend in various complicated ways on the attitudes of individuals who make it up, those attitudes will often come apart from the attitudes of individual judges, jurors and jailers.[23]

So advocates of UCE should take the subject of the expressive attitudes which are constitutive of punishment to be the state. And advocates of VRE, who take those attitudes to include the attitude of vindictive resentment, should take the subject of that attitude to be the state as well. And now we do have a situation where we seem compelled to see punishment and reconciliation as involving conflicting attitudes within the same subject. For the arguments of Section IV suggested that in the kinds of case we are interested in – those involving large-scale wrongdoing on the part of the state – the attitudes characteristic of reconciliation would themselves need to be attitudes expressed by the state.

Can we avoid the tension here by suggesting that the apparently conflicting attitudes involved in punishment and reconciliation can be ascribed to the same subject at different times? We might think that when matters are in good normative order, punishment (with its associated attitudes) will come first, and reconciliation (with its associated attitudes) afterwards. If so, the fact that the attitudes involved are rationally incompatible need raise no problems. However this seems an unsatisfactory way of resolving the tensions between punishment and reconciliation. If we go no further than this we seem compelled to accept that while the process of bringing wrongdoers to trial is still ongoing, reconciliation cannot start; and it also suggests that at a certain point this process must come to a permanent stop in order for reconciliation to start. Given normal human limitations on the speed at which trials can take place, this in turn

[21] (An advocate of VRE will of course agree with this. This is unsurprising since VRE is simply a way of adding to UCE some of the specificity which UCE by itself lacks

[22] For conceptions of the state along these lines see Gilbert 2008; Stilz 2009, and for an application to the philosophy of law see Shapiro 2011. For a more detailed development and defense of the idea that it is the form of expression of a collective agent that we should be considering here which is in question here see Wringe 2016 chapter 3.

[23] This is not to say that the state's attitudes are do not depend in any way on the attitudes of its members, but only that it may have attitudes that do not match those of any given set of individuals. See List and Pettit 2010 for a defense of the idea that the attitudes of group or collective agents cannot be reduced to the attitudes of individuals

suggests that if reconciliation is to take place, may perpetrators will come to benefit from either an official or de facto amnesty for their crimes. We might hope to resist this conclusion. If so, it seems as though the suggestion that the attitudes involved in punishment and reconciliation belong to the same subject at different time is not enough to resolve our difficulties.

In this Section I have argued that punishment and reconciliation appear to involve conflicting attitudes on the part of the state. This should be enough vindicate the claim that I made in Section I that there is indeed a tension between punishment and reconciliation. However, I think that the idea that the punishment and reconciliation involve opposed attitudes in distinct subjects is problematic for a second reason as well; one which turns on questions of legitimacy as well as logic. The suggestion depends on the idea that the attitudes expressed by the state might be radically distinct from those belonging to the citizens. I have argued, in ways which are echoed by much of the literature on collective action and collective agency that there is no logical or metaphysical; difficulty with this view. Nevertheless, we might hope that, in democratic societies at least, there would be some kind of congruence between the attitudes expressed by a state and the attitudes expressed by the citizens of a state.

It is natural to think of the state and its institution as speaking for the citizens as a collective body. To the extent that those institutions give voice to sentiments which its citizens are unable to endorse, it fails in one significant purpose.[24] Indeed, in a well-functioning democratic state we can expect a significant number of citizens to endorse and identify the views expressed.[25] We might even think that the existence of a significant conflict between the attitudes expressed by the citizens and those expressed by the state counts against the legitimacy of the state. In other words, although the attitudes constitutively involved in punishment and reconciliation can belong to different subjects, in legitimate states at least some of the attitudes involved in punishment will also be held by individual citizens. So again, we have conflicting attitudes in a single subject: not a collective subject this time, but an individual citizen.

We should also notice that a situation where individual citizens are supposedly reconciled to one another but where they collectively give voice to a form of vindictive resentment seems unlikely to be stable. For it seems unlikely that those who are on the receiving end of the state's vindictive resentment will refrain from viewing it as an expression of the attitudes of those citizens of the state who happen to have the upper hand politically. So a solution of this sort may also turn out to be unsatisfactory from this point of view. (One might object that this final consideration is psychological rather than political. But I do not think it is unreasonable for a philosopher who is concerned with political issues to be concerned with the long term stability of possible solutions to

[24] Gilbert 2006

[25] There are other problems with the idea that political reconciliation should simply be seen as a relationship between citizens. In many cases, the state is one of the agents which has been implicated in harms to citizens. To see it as standing outside the population of individuals between whom reconciliation is required seems to overlook something important. Furthermore the suggestion seems to involve a conception of reconciliation which at least some theorists of political reconciliation would regard as unduly limited. On views like this political reconciliation necessarily has a public dimension to it. Thus for example, Colleen Murphy (Murphy 2010) has argued for a view of political reconciliation on which reconciliation constitutively involves the rule of law. This can't simply be something constituted by the attitudes of individuals: it must be realized in public institutions.

problems with a practical dimension. There is certainly eminent precedent for such concern, among philosophers otherwise as different as Plato and Rawls.)

6 The Expressive Strategy

In explaining the tension between punishment and reconciliation, I appealed to Feinberg's view that punishment constitutively involves the expression of vindictive resentment. One reaction might be to abandon the idea that punishment has any kind of expressive dimension. However, I argued in Section II that those who take punishment to be capable of being justified should accept some form of expressivism, in the shape of UCE. So we have reasons to consider other ways of responding to the tension. [26] Paying attention to the exact form that expressivism takes – and in particular to exactly which attitudes punishment is supposed to express will prove to be a more promising strategy.

Feinberg held that the expressive role of punishment raised difficulties for the philosophical project of justifying punishment.[27] More recent authors have, however, attempted to appeal to the expressive role of punishment as a way of justifying the institution, or at any rate of undermining arguments against its justifiability. In particular Antony Duff and Christopher Bennett have appealed to the expressive role of punishment in order to respond to two significant critiques of punishment.[28]

Duff argues that since punishment should be conceived of as a form of communication which is aimed at evoking regret or remorse in a convicted offender, it can legitimately be regarded as treating offenders as autonomous sources of agency: punishment involves a form of communication which aims at evoking a particular kind of response in an offender, but the offender can decide whether or not to respond in the way sought.[29] By contrast Christopher Bennett develops the idea that punishment has an expressive role in a rather different way, in order to respond to a slightly different critique.[30] Bennett's concern is with the suggestion – prominent among advocates of 'restorative justice' - that criminal punishment involves the state inappropriately inserting itself into and taking possession of a dispute between two individuals in ways which can frustrate, rather than promote their efforts to live as members of a political community. [31] He argues that we should see the institution of punishment being something which enables a convicted offender to express a form of symbolic penance of a sort which might enable him/her to be reintegrated into the community.

These two accounts appeal to very different ways of developing expressivist insights in response to moral critiques of punishment. However the insights which underlie them need not be incompatible with one another. We might conceivably regard

[26] Feinberg's version of expressivism was first put forward in 1965. Expressivism has been defended by Duff et al. (2007, 2001); Hampton (1992); Metz (2000, 2007); Lippke (2007); Bennett (2008); Glasgow (2015) and Lee (2015). Many of these views conflict with Feinberg on a number of issues, including the question of what exactly punishment expresses and who it expresses it to, For my own views see Section IX and Wringe 2016a chapters 2–4

[27] Feinberg 1965

[28] Duff 2001; Bennett 2008

[29] Duff 2001. The critique he is addressing originates with Morris 1968

[30] Bennett 2008

[31] Bennett 2008 pp 13–26. The critique to which he is responding originates with Christie 1976

punishment as having more than one expressive function. In particular there seems no reason to think that punishment could not function so as to allow communication from a society to an offender against whose rules he or she has transgressed, and a form of communication from the victim to the offender.

Both expressive accounts seem to offer scope for dissolving apparent tensions between the aims of criminal punishment for wrongdoing and reconciliation. First consider Duff's account, on which criminal punishment is aimed at evoking remorse and regret for wrongdoing. Duff suggests that experiencing these emotions may play an important role in allowing a wrongdoer to become reintegrated into a political community.[32] We can spell out why by appealing to a plausible thought about the relationship between remorse and forgiveness. For we might hold that forgiveness is an appropriate response to wrongdoing when the individual to be forgiven shows some sort of remorse for their wrongdoing. If forgiveness of wrongdoing plays a central role in political reconciliation then the idea that the goal of criminal punishment can constitutively contribute to reconciliation, rather than standing in tension with it seems vindicated.

However Bennett has argued that even if we regard a wrongdoer's remorse for wrongdoing as a potentially valuable result of punishment, it ought to be no business of a liberal state. Liberal states should not – Bennett suggests - try to 'get inside the head' of offenders in the way which Duff appears to advocate.[33] Whether or not Bennett is correct about this, an analogous thought to his appears to carry some weight here. Even if we accept that state punishment of criminal action aims at bringing about remorse and regret, we should also accept that in states which have been marked by large scale public wrongdoing its right to do so will be contested by many of those who undergo punishment. We might expect that a penal regime that aims at these goals will in practice result not in reconciliation but in forms of resentment which run the risk of undermining that goal both on their own and by giving citizens who have been wronged good reasons to avoid regarding any process of reconciliation as having taken place.

Fortunately, Bennett's version of expressivism seems to provide us with the means to respond to this criticism. For on his view, what is significant about punishment is not the emotions that it aims at evoking, but the fact it provides the means for a public expression appropriate to those emotions, whether or not the offender feels them.[34] Bennett's idea seems to be that that this symbolic expression puts a victim in a position to accept that an offender has made amends, even where the offender does not have the kinds of mental attitude that would most appropriately be expressed in this way.

The line of argument outlined here relies on the idea that there are close connections between remorse and the appropriateness of forgiveness and between forgiveness and reconciliation. However, as we shall see in the next section, it is harder to spell out these connections in a satisfactory manner than we might initially suppose.[35]

[32] Duff 2001 pp 42–8
[33] Bennett 2008.
[34] Bennett 2008 pp 125ff
[35] Benn 1996; Govier 1999; Hieronymi 2001

7 Reconciliation and Forgiveness

The idea that reconciliation either necessarily or typically involves forgiveness may initially seem plausible.[36] When we think of two individuals becoming reconciled in a situation that has involved some form of prior wrongdoing on the part of one or both of them, we typically think that forgiveness of one individual by the other or of both individuals by each other as plays either a causal or a constitutive role. However, as I noted in the opening section of this paper, the notion of reconciliation that we ought to be concerned with in this context is political, rather than interpersonal reconciliation. There are several reasons for being doubtful as to how closely we can model the notion of political reconciliation on that of personal reconciliation,

First, as Susan Dwyer notes, incorporating the notion of forgiveness into our conception of political reconciliation runs the risk of associating it too closely with Christian religious traditions which, while powerful and resonant in some societies which stand in need of reconciliation, may be out of place and seen as an alien imposition in others.[37] Secondly, forgiveness seems naturally associated with the idea of returning a disrupted relationship to its previous state. However, as many authors have pointed out, in many states which might reasonably aim at political reconciliation, there is no past period of morally acceptable relationships which forgiveness might aim at restoring.[38] Thirdly, political reconciliation seems to be a relationship which holds between groups, whereas forgiveness is, at least paradigmatically, a relationship between individuals.

We might attempt to map the relationship of reconciliation between groups on to that of forgiveness in two different ways. One would be to suggest that groups are reconciled when every individual in each group has forgiven every individual (or at least every individual wrongdoer) in the other. This seems problematic for two reasons. First, it seems to set the bar for reconciliation impossibly high. It thereby fails one of the conditions on a conception of political reconciliation which I identified at the beginning of the paper. Secondly it seems to ignore the fact that there is a public political dimension to reconciliation which seems not to be a part of individual acts of forgiveness, no matter how many of them we aggregate. Individual acts of forgiveness need have no such public and political dimension.

An alternative way of mapping the relationship of reconciliation on to that of forgiveness would involve taking seriously the idea of groups forgiving one another. Several authors have addressed the question of whether it makes sense to think of one group forgiving another. It raises at least two distinct issues: whether groups are the kinds of things which can forgive; and whether they are the kinds of things which can be forgiven. Trudy Govier has explored the first question; Linda Radzik and Charles Griswold the second.[39] They have all given affirmative answers to the questions they have considered. However, their discussions have focused on groups which are collective agents: states, governments, universities religious orders and the like.[40] Groups of

[36] For some skepticism on this score see McNaughton and Garrard 2010 pp 57–62
[37] Dwyer 1999
[38] Moellendorf 2007
[39] Govier 2002; Griswold 2007; Radzik 2009
[40] Govier 2002 pp 87–8; Griswold 2007 pp 147–157; Radzik 2009 pp 175ff

this sort typically have a complex internal structure which makes them capable of agency.

Situations which call for political reconciliation will often involve groups which are collective agents as possible objects and subjects of forgiveness: one need only think of the role of paramilitary organizations and the RUC in Northern Ireland, the ANC and the South African security forces in South Africa and so on. However, these are not the only kinds of groups whose relationships are relevant to political reconciliation: in many cases reconciliation (if it takes place) will be a relationship between relatively unstructured groups, such as ethnic and cultural groups, which do not meet the conditions for collective agency.

It is much less clear that groups of this sort can be subjects of forgiveness. Some accounts, = such as Margaret Gilbert's account of plural subjecthood = seem to leave open the possibility that they might be able to. However, it is not clear that it would be helpful to incorporate a notion of group forgiveness built on Gilbert's account into an account of political reconciliation. There are two reasons why not. First Gilbert's conditions for a group being in a collective mental state seem require group members to identify themselves with that a group. We might think that doing so, they do something that makes it less likely that members of different groups will see one another simply as citizens. But we might think that in many cases political reconciliation will be better promoted by encouraging citizens to identify themselves as citizens rather than as members of this or that particular group.

Furthermore, if we adopt a model of intergroup forgiveness modeled on Gilbert's account we are likely to run into a problem analogous to the problem we faced when we considered the possibility that reconciliation might require every member of a group to forgive every other member of a group. Gilbert's account of the conditions under which a collective subject can be held to be the subject of an attitude require that each member of the group be committed to holding a collective attitude as a body. [41] Where large unstructured groups are concerned, this condition seems hard to meet. There is obviously no logical objection to requiring that every member of the group should hold this commitment. Ensuring it in practice is another matter. If we adopt an account of reconciliation on which group forgiveness, understood along the lines suggested by Gilbert's theory, is a necessary condition for reconciliation, then reconciliation will turn out to be an unachievable political goal.

8 Political Reconciliation Without Forgiveness

Although the arguments of Section VII do not show conclusively that there could not be a workable notion of political reconciliation in which forgiveness plays a central role, they present serious obstacles to elaborating such a conception. A different approach may be more fruitful. In this section I consider two accounts of political reconciliation in which the notion of forgiveness does not play a key role: those put forward by Colleen Murphy and Darrell Moellendorf. [42] I argue that Moellendorf's account is preferrable to Murphy's on the grounds that it does justice to the intuitively

[41] Gilbert 1989, 2001

[42] Murphy 2010; Moellendorf 2007

plausible idea that reconciliation needs to have a backward looking dimension in a way that Murphy's does not. (There might conceivably be other views on which something other than political regret plays the backward-looking role which political regret plays on Moellendorf's view. But I am not aware of any such views. So on these admittedly non-conclusive grounds, I recommend the adoption of Moellendorf's view.)

For Murphy, reconciliation is a process which aims at respect for moral agency and reciprocity in a society which has been marked by an absence of these features in its political life. She suggests that these features of political life are best ensured by a commitment to (and actualization of) the rule of law; by action which creates and sustains political trust, and support for key capabilities of agents.[43] This conceptions of reconciliation involve thinking of reconciliation as involving a state of society and of the political order which may depend, causally, on certain kinds of interpersonal relationship between individuals, but which is not constituted by them. The extent to which forgiveness of wrongdoing will be either a necessary condition of reconciliation, or a factor contributing to its realization remains open.

On Murphy's conception of reconciliation, criminal punishment can contribute to the goal of reconciliation in at least one way. For as she argues, such processes may be seem as both sustaining and embodying a commitment to the rule of law. However there are several reasons for thinking that her account does not go far enough in showing that the goals of criminal punishment and reconciliation do not stand in tension with one another.

First, as Fletcher and Weinstein have argued, it seems plausible that however we conceive of reconciliation, successful processes of reconciliation in heavily divided societies will require attitudinal change among members of groups who have in one way or another benefited from oppression, but may not become the focus of criminal prosecution.[44] Fletcher and Weinstein make a convincing case that the individualizing focus of criminal trials can undermine the chances of attitudinal change of this sort. Secondly, Murphy and several other authors take political trust to be a necessary component of political reconciliation. However, the use of criminal trials to achieve what look like political goals seems to create fertile ground for the growth of cynicism among members of groups who are regarded as responsible for oppression. Finally, respect for the rule of law may itself limit the role which criminal trials can play in the process of political reconciliation, especially to the extent that prosecution of oppressors is seen as involving either retrospective legislation or something akin to Victors' Justice.

Murphy's attempt to dissociate the notion of reconciliation from that of personal forgiveness and her focus on the social relations which ought to provide in a society which has achieved reconciliation provide a valuable change of focus from notions of reconciliation which are centered on the idea of personal forgiveness. Nevertheless, her view seems somewhat counter-intuitive. It seems to neglect what one might call the 'backward-looking' dimension of reconciliation. On the view that she puts forward it seems that a society, might count as having achieved reconciliation without any form of widespread acknowledgment that the previous state of a society was unjust and oppressive. But a state of affairs of this sort seems more justly described as involving amnesia rather than reconciliation.

[43] Murphy 2010
[44] Fletcher and Weinstein 2002

Darrell Moellendorf has suggested that a plausible conception of reconciliation must involve what he calls 'political regret'. He writes:

> The establishment of a peaceful political community after a period of conflict arising from severe injustice requires resolution of some of the central terms of the conflict; and acceptance of the terms of the resolution often requires public expressions of appropriate' regret about at least some of the injustices that occurred. Once the grounds for regret are clear, public policy may express this in a number of ways including engaging in symbolic acts of reparation and commemoration, instituting name changes of public buildings or streets, and establishing public holidays. These are properly understood as, at least in part, expressions of regret because they acknowledge that the injustices ought not to have happened. [45]

Moellendorf refers to the kinds of public expressions of appropriate regret as expressing 'political regret. A conception of reconciliation which incorporates the expression of political regret will have the kind of backward-looking dimension that Murphy's conception lacks.

Moellendorf expresses his point in a way which suggests that political regret is a causal precondition of the kinds of social relations which he takes to be constitutive of reconciliation. Put in this way, the point is one which Murphy might be able to accommodate: she might, perhaps, argue that this kind of public acknowledgment of wrongdoing is an important causal precondition of the kinds of action needed for establishing public trust.

However, this seems to be an unsatisfactory way of acknowledging the significance of Moellendorf's point. For it seems plausible that public acknowledgment is something which those who have been oppressed might be entitled to demand as a matter of right, even if as a matter of fact society could continue without making such an acknowledgment. If we think that the only role which public acknowledgment of this sort plays in reconciliation is that it is an instrumentally valuable precondition of other kinds of social arrangement, then it is hard to see how this could be true: it seems as though a society that managed to find a way of carrying on without making public acknowledgments of wrongdoing would have no reconciliation-related reason to do so. [46]

Moellendorf's point might be understood in a somewhat different way, however. We might regard the expression of political regret not merely as a causal precondition of reconciliation, but partially constitutive of it. The examples which Moellendorf gives of ways in which political regret can be expressed might make this suggestion seem unpromising. No-one could regard it as *constitutive of* reconciliation in an interesting sense that it should involve the institution of public holidays or the renaming of public buildings – reconciliation would presumably be conceivable in a society which had neither public buildings nor public holidays. But it is consistent with this to suggest that it might be constitutive of reconciliation that it should involve the public acknowledgment of wrongdoing in some respect, and in ways which allow for the possibility that

[45] Moellendorf 2007

[46] I am grateful to an anonymous reviewer for *Philosophia* for making clear the need to spell out the argument for rejecting this interpretation of Moellendorf's position.

this acknowledgment can be demanded as of right. But to say this is to suggest that more may be required for the expression of public regret than the renaming of buildings.

Why should we take the notion of political regret to play a central role in political reconcilation? And why we should talk of specifically political regret? To address these questions it will be helpful to engage in some analysis of the idea of regret in general and some exegesis of Moellendorf's notion of political regret in particular. I take regret to be something which looks at some past state of the world, acknowledges some form of agential involvement in it on the part of the person doing the regretting, and acknowledges that it was bad. Political regret is a more complex notion. One idea which seems to be central to the notion of political regret which Moellendorf is attempting to articulate here, is that it involves the *public expression* of regret. *T*wo aspects of publicity which seem to be important here. The first is that the expression of regret is made *on behalf of* the public. The second is that it is made *to the public*.

So Moellendorf's claim is that reconciliation involves the public acknowledgment of wrongdoing. We might wonder why this should be so. Couldn't there be political reconciliation without any acknowledgment of wrongdoing at all? And couldn't there be one in which there was some acknowledgment of wrongdoing, but only on the part of private individuals, or to private individuals? These possibilities seem harder to substantiate than one might initially imagine. Let us start by considering what might enable us to distinguish a society in which reconciliation had taken place from one which suffered from a condition of political amnesia. One thing that would seem missing in such society is a sense that its members could engage in common enterprises in which the full humanity of everyone involved was acknowledged. It seems difficult to see how such acknowledgement could be made without some form of repudiation of prior acts which appear to deny that humanity. And it hard to see what from such repudiation could take other than expressions of regret.

We might still wonder why there should be any need for this regret to be publicly expressed. But here some of the considerations to which I appealed to in section II when explaining why we might be interested in a specifically political conception of reconciliation at all seem relevant. But there might be another point which is worth making. A society in which reconciliation has occurred is presumably not one in which people need to be able to engage in any kind of common enterprise, but one in which they need to be capable of doing so in specifically public contexts: as members of parliament, judges, jurors, members of school boards and so on. For these purposes, it seems as though repudiation of past wrongdoing taking place between private individuals may not be enough. For while this might be enough to ensure that individuals saw one another as equals in private life it would not be enough to enable them to be sure they were seen as equals within the political domain. For this, something more seems to be needed. Political regret, in so far as it involves a repudiation of prior wrongdoing to the public and on behalf of the public seems to fill the gap. [47]

One might doubt whether Moellendorf's conception of political reconciliation really is a conception of reconciliation as that notion is typically understood. For we might think that a proper understanding of the notion of reconciliation is one which takes it to

[47] It's worth noting that although I claim political regret is a necessary condition for the right kinds of relationships between citizens in the political domain, I don't claim that it is a sufficient condition.

be a relationship that can hold (or fail to hold) between the members of a society.[48] It might seem as though, on Moellendorf's view, political reconciliation is a relationship which holds (or fails to hold) between a state and its citizens. For, insofar as public acts are required for political reconciliation, these seem to be acts performed by the state or the public. However, this objection misses something significant about the relationship which can hold between acts of the state and individuals in a well-functioning state. For the expressive actions of such a state can appropriately be seen as expressive actions taken on behalf of individual citizens that make it up. If so, then political reconciliation of the sort that Moellendorf envisages can properly be understood as a relationship between citizens, even though its existence depends on actions taken by the state. [49]

9 Expressivism Revisited

In the remainder of this paper, I shall suggest that the punishment of wrongdoing may play a significant role in the process of reconciliation by being a vehicle for the public expression of regret. In doing so, I shall draw on a way of understanding the expressive view of punishment which is rather different from the one put forward by either Duff or Bennett – one which emphasizes what Uma Narayan calls the 'denunciatory' as opposed to the 'expressive' dimension of punishment.

Denunciatory theorists of punishment differ from communicative theorists on the issue of who the intended audience of the expressive acts that are constitutive of punishment is. For communicative theorists, like Duff the audience is (at least in the first instance) an offender. An alternative view – the denunciatory view - is that the relevant audience is the members of the political community whose rules have been infringed.[50] In typical cases, the offender and the victim of the crime (if there is one) will both be members of this community.

Those who are committed to UCE may accept a denunciatory account of punishment, but need not do so. And denunciatory theorists may, but need not, accept VRE. To adopt a denunciatory theory is to take a view about the intended audience of punishment, not about its content. A denunciatory theorist might reject VRE by suggesting that what punishment expresses is a claim that certain norms are in force, and are expected to be taken seriously by members of the political community. This view of punishment does not account for the phenomenology of the experience of being in prison which, as we saw in Section III, Feinberg cites in support of VRE. However we might respond by saying that there is in any case something pathological about this phenomenology in any case, and that a humane prison system would seek to reduce it.

By contrast, this form of denunciatory expressivism does seem to be capable of explaining the secondary symbolic functions of punishment to which Feinberg appeals in defending VRE. Those functions were the authoritative disavowal of certain forms of

[48] Or perhaps between different groups in a society. See footnote 18 for further discussion of this suggestion.
[49] I am grateful to an objection from an anonymous reviewer for *Philosophia* for a question which prompted this clarification.
[50] This terminology is from Narayan 1993. Authors whose views have a denunciatory element include Primoratz 1989; Hampton 1992; Metz 2000, 2007. For a detailed defence of a denunciatory view see Wringe 2016a, Wringe 2016b.

behaviour; symbolic non-acquiescence in that behaviour; vindication of the law; and absolution of non-criminals. As I observed in Section III, none of these functions of punishment seem to require that what is expressed in punishment be vindictive resentment. I here note that if punishment that expresses claim that certain norms are in force, and are expected to be taken seriously by members of the political community, then it seems capable of fulfilling these further symbolic functions.

I argued in Section II that we should accept UCE. I have. also argued that someone who accepts UCE may accept a denunciatory version of expressivism. But is there any reason why they should? The denunciatory view of punishment is a view about the audience of punishment. So we should first ask why we should take the expression which, according to UCE, is constitutive of punishment to have any audience at all, and secondly why we should take this audience to be the political community in particular.

Some versions of expressivism are what I have elsewhere called 'audience-independent'.[51] On this view, punishment must have an expressive component; but the expression which is constitutive of punishment need not be directed at anyone at all. This is a somewhat strange view. If it was correct, it would be hard to see why we should take the expressive component of punishment to be significant or interesting. Furthermore it is worth noting that on a view of this sort, unlike the denunciatory view, the expressive function of punishment could not play a role in allowing punishment to fulfill what Feinberg calls its secondary symbolic functions. For those functions depend on communication with an audience.

If we set audience-independent views aside, then we need to ask who is the audience of the expression involved on in punishment might be. The most plausible answers to this question are that it is the offender (as communicative theorists such as Duff claim) or the political community at large (as I claim).[52] I shall argue that the communicative view has severe drawbacks and that we should therefore prefer the denunciatory view.

One reason is that it is unclear how or whether Duff's communicative account of punishment can accommodate certain non-paradigmatic instances of criminal punishment, such as the punishment of corporate entities and of war criminals. In both cases, though for slightly different reasons, the goals of communicating disapproval for the sake of evoking regret or remorse seems out of place. Where corporate wrongdoing is concerned this is because, even if it is possible to make sense of a corporate entity experiencing remorse, it is implausible that the remorse of a corporate agent is valuable or significant in the same ways or for the same reasons that the remorse of an individual might be regarded as valuable or significant. In particular it seems implausible to regard corporate entities as members or potential members of a political community in the same way that individual agents are.[53] In the case of war criminals, it is because it seems counter-intuitive to suppose that the most appropriate way to respect their autonomy is to offer them the chance of integration into communities of which they may never have been nor wished to be members.[54] By contrast a denunciatory account seems to be able to handle both kinds of case with comparative ease.

[51] Eg Glasgow 2015. For further discussion of such views see Wringe 2016b

[52] Other possibilities are logically possible: we might take the intended audience of legal punishment to be the officials of the legal system; or God; or posterity. But as far as I know, no advocate of an expressive view has defended any of these possibilities. So I shall not consider them here.

[53] Wringe 2012

[54] Wringe 2006

It is unclear quite how much an advocate of a communicative view need concede in response to cases like this. One might be tempted by an exclusively denunciatory view of the expressive role of punishment. One apparent obstacle in the way of such a view would be Duff's view that it is only the communicative aspects of punishment that ensure that punitive treatment can succeeds in treating offenders as autonomous.[55] An alternative would be to retain a communicative conception of the punishment of individuals and resort to a denunciatory conception only for non-paradigmatic cases. Call this a 'disjunctive hybrid view'.

The disjunctive hybrid view has significant drawbacks. One is that it suggests that the non-paradigmatic instances of punishment only count as instances of punishment in a loose or extended sense. Furthermore, it ought to seem problematic to anyone who is tempted to think – as many advocates of communicative views are - that procedural aspects of the criminal trial – such as the presumption of innocence and the high burden of proof can be accounted for and justified by reference to communicative considerations.[56] For these features seems to be equally desirable in the kinds of non-paradigmatic cases of punishment which we have been considering. It is hard to see why the norms relating to two different forms of expression should justify similar procedural features in the two kinds of case.

An alternative approach would involve seeing paradigmatic instances of punishment as having two communicative functions: one denunciatory, and one communicative. Call a view of this sort a 'Conjunctive Hybrid View'. This view seems well able to account for a further aspect of the criminal trial which might otherwise seem hard to explain: namely, its publicity. In recent work, Antony Duff and Victor Tadros have made much of the idea that crimes are 'public wrongs' in the sense that they are wrongs with which the public is properly concerned.[57] However, it does not follow from this conception of publicity that trials and punishment should be public in a second sense – namely that of being open to scrutiny by the public.

If we think of the punishment of criminals as involving a form of communication which is directed, in the first instance at the offender it seems mysterious why trials and punishment should be public in this sense. Indeed there are at least two reasons for thinking that it should not be. The first concerns the effectiveness of the communication. We should at least be open to the possibility that private communication will be more successful than communication which is open to public view. The second reason has to do with the rights of the accused. Insofar as exposure to public gaze might itself be taken as a form of harsh treatment we might wonder how it is to be justified. However, if we think that the denunciatory aspect of punishment is either constitutive of punishment, or plays a central part in explaining how punishment can be justified, the public nature of the trial will no longer seem mysterious. For on this account, it will be central to a form of harsh treatment's constituting punishment that it involve a communication with the public. But this will be ensured by the trials and imposition of punishment being public in the sense we have been considering here.

[55] For some skepticism on this score see Wringe 2006
[56] Duff 1986
[57] Duff 2001; Tadros 2005

10 Public Regret and Denunciatory Punishment

In Section VII I argued that conceptions of reconciliation which incorporated either personal or group forgiveness were so demanding as to be unacceptable as conceptions of political reconciliation. I then argued in Section VIII that a conception of political reconciliation which did not incorporate forgiveness should nevertheless incorporate some backward-looking aspects, and I suggested that we could meet this constraint if we took the expression of political regret to be partly constitutive of reconciliation. I now wish to conclude by arguing that the denunciatory role of punishment, for which I argued in Section IX, makes it a suitable means for expressing political regret in situations where that is what political reconciliation requires.

In order to pre-empt misunderstanding, it will be helpful to limit what I am claiming in two respects.[58] First, I do not want to claim that criminal punishment should be the only way in which political regret is expressed in societies which are aiming at reconciliation. More importantly. I do not want to claim that its pre-eminence is such that it follows that political regret can only appropriately be expressed by pursuing each and every instance of wrongdoing. Often this will be counter-productive: as the South African example shows, there will be occasions on which amnesties for certain kinds of wrongdoer (or certain kinds of wrongdoing) will contribute more to reconciliation than over-zealous prosecution. However, it will follow from what I say that easy resort to amnesty will undermine the possibility of political reconciliation.

As we saw in Section VII, a key part of Moellendorf's notion of political (as opposed to personal) regret is that it involves the public acknowledgment of wrongdoing.[59] This acknowledgment needs to be public in two senses which will already be familiar: it should expressed in public, and it should be expressed by someone, or some institution, which has the right to speak on behalf of the public. It will also often need to be public in a third sense: it will need to acknowledge that the wrongdoing that occurred was committed on behalf of – or was sanctioned by – an institution which took itself to speak on behalf of the public.[60] It needs to be backward-looking in acknowledging wrongdoing, but as Moellendorf emphasizes, it also has a forward-

[58] It is also worth forestalling a further misunderstanding of my view, which occurred to an ingenious anonymous reviewer for *Philosophia*. My claim is that the form of denunciatory expression which is constitutive of punishment can - under the right circumstances - be a way of expressing political regret. For punishment to be a way of expressing political regret, certain further conditions must be satisfied. For example, only some kinds of actions can be the object of political regret. To a first approximation, a crime can only be the object of political regret if it is either committed by or condoned by a body with political authority. So for example many crimes committed by individuals for their own purposes could not be the object of political regret, since they do not have any political dimension. Since there is no such thing as political regret for such crimes, there is no such thing as expressing political regret for them. So punishment will not express political regret for them. As a result, I do not hold that punishment must *always* express political regret. So it would be a serious misunderstanding of my view to take it to entail that actions performed by the state in response to individual crimes such as petty thievery cannot be punishment, because they do not express - or cannot plausibly be taken to express - political regret.

[59] Moellendorf 2007

[60] This third condition – that punishment is for acts carried out on behalf of, or sanctioned by the state = explains why not all instances of punishment will express political regret. (As I noted in Section VI, one cannot regret, in the relevant sense, actions that are not in any way one's own.)

And this is what we should expect, since reconciliation and political regret are only appropriate in some states, some of the time.

looking dimension, since regret involves 'a commitment to preventing similar injustices in the future.'

Can the punishment of wrongdoing express political regret? If the denunciatory conception of punishment is correct, it seems as though it can. On that view, punishment can be regarded as involving the expression of attitudes by an institution which has the right to speak on behalf of the public if any institution does: namely, the state. It will also involve the acknowledgment of wrongdoing – and, insofar as punishment is inflicted in accordance with settled law – some kind of commitment to similar wrongdoing being prevented in the future. It is also public in at least one of the senses of publicity which I have isolated here: it is a commitment which is made to the public and in full view of the public.

To say that punishment can express political regret is not to say that it must always do so. So it would be a serious misunderstanding of my view to take it to entail that actions performed by the state in response to individual crimes such as petty thievery cannot be punishment, because they cannot plausibly be taken to express - political regret.[61] A crime can only be the object of political regret if it is either committed by or condoned by a body with political authority. Many crimes committed by individuals for their own purposes, do not fall into this category. Such crimes cannot be the object of political regret. Since there is no such thing as political regret for such crimes, there is no such thing as expressing political regret for them. So punishment cannot express political regret for them.

The idea that political regret should be expressed by a body which has the right to speak on behalf of the public places some significant limitations on the kinds o punishment which might appropriately be regarded as expressions of public regret. For example, we might doubt that vigilante justice could play this role. Perhaps more controversially, it is also unclear whether punishments imposed by bodies external to the state such as international tribunals could play this role. For it is at best unclear whether they could be regarded as speaking for the public in the relevant sense.[62]

Are matters are different when we consider the state as an agent of punishment? I claimed in section V that the expressivist should think of the attitudes expressed in punishment as being attitudes expressed by the state; and it is central to the denunciatory version of expressivism that I have been putting forward here that this should be so. Indeed given the role which the legal system plays in regulating the everyday life of citizens, the legal system seems especially well-suited to express the public acknowledgment of wrongdoing.

Nevertheless, some might wonder whether we can regard the state as speaking on behalf of the public in societies which are deeply divided in the way which we might expect societies marked by large-scale public wrongdoing to be. This objection is far-reaching: if sustained, it seems to apply not only to the judicial system, but to any other state institution. To that extent it seems to be an objection not simply to my view of the

[61] I am grateful to an anonymous referee for Philosophia for giving me the opportunity to address this ingenious misunderstanding of my view.

[62] I here make good the promissory note issued in footnote 2. I should note that I do not want to rule out the possibility that international organizations might speak for the public on at least some occasions. Exploring the conditions under which they might do is a large job which would take me beyond the scope of this paper, but for some preliminary considerations see Wringe 2016a chapters 7 and 8.

role punishment can play in reconciliation, but to the conception of political reconciliation developed in Section VIII.

My response is to suggest that the claim of the law to speak on behalf of the public is always to some extent both provisional - since existing laws can be undone by the public through the legislative process - and aspirational = in the sense that the political community on behalf of which the law purports to speak is one which is also partially constituted by the law which it (the political community) may choose to abrogate. If so, what we have in the case we are considering is merely an extreme case of a much more widespread phenomenon.

Finally, we need to consider a further sense of publicity. One might think that for regret to be public in the sense required by Moellendorf's analysis, it must be regret which is not only expressed by the public, but regret whose object or target is acts carried out by the public. Someone might doubt whether judicial punishment can properly express regret which is public in this sense for the following reason. If we rule out the possibility of retroactive legislation – something which seems incompatible with any conception of reconciliation which incorporates the idea of the rule of law, the punishments meted out by the courts will need to be punishments for actions which were against the law at the time they were committed.

However, on certain conceptions of what it is for a state to act, such actions cannot be conceived of as actions of the state. For it might be argued that something can only count as the act of a collective body if it is carried out in accordance with a certain 'corporate decision procedure', and that in the case of states, the law is partly constitutive of the corporate decision procedure for the kind of corporate body that a state is.

I think that this line of thought is misconceived: it involves thinking of the state as a collective agent in the wrong kind of way. The notion of a collective decision procedure seems to be a notion that is best defined in legal terms, If so, and if we also think of the law as embodying some form of collective agency (as Scott Shapiro and Margaret Gilbert have recently suggested) we will have to think of there being a prior form of collective agency involved in instituting the state.[63] This leaves some room for the idea of there being actions which are carried out by the state even though they stand in opposition to the state's explicit declarations. Of course, this may require us to regard certain some kinds of official declarations as being sincere even though they were never intended as such, or of carrying a kind of meaning they were never intended to have. But if that is something which is required for reconciliation to be a possibility in a divided community, it seems as though it is something which might be gladly embraced.

Acknowledgements Earlier versions of this paper were given at a workshop on Punishment, Forgiveness and Reconciliation at the MANCEPT workshops in Political Theory in September 2014 and at Forgivness and Reconciliation VIII in Dubrovnik in May 2015.I am grateful to the organisers of those workshops for the opportunity to present my work, to the audiences on those occasions for feedback and to Monica Mookherjee and William von Bulow for encouraging me to present my work in these venues. I'm especially grateful to Paula Satne, both for organising the MANCEPT workshop, and for her encouraging remarks on a number of earlier versions of the paper, without which I would have been lost. I also received valuable written comments from Jack Woods, Sandrine Berges, Katie Stockdale, Fragano Ledgister, David McNaughton, and Janet Brennan Croft. I should also mention the contributions of an anonymous referee for this journal whose comments on a number of ealier versions strongly influenced the content and form of the final version of the paper.

[63] Gilbert 2006; Shapiro 2011

References

Benn, P. (1996). Forgiveness and loyalty. *Philosophy, 71*, 369–384.

Bennett, C. (2008). *The apology ritual*. Cambridge: Cambridge University Press.

Boonin, D. (2008). *The problem of punishment*. Cambridge: Cambridge University Press.

Duff, R. A. (2001). *Punishment communication and community*. Oxford: Oxford University Press.

Duff, R. A. (2007). *Answering for crime*. Oxford: Hart Publishing.

Duff, R. A., Farmer, L., Marshall, S., & Tadros, V. (2007). *The trial on trial vol 3: Towards a normative theory of the criminal trial*. Oxford: Hart Publishing.

Dwyer, S. (1999). Reconciliation for realists. *Ethics and International Affairs, 13*(1), 81–98.

Falls, M. (1987). Retribution, reciprocity, and respect for persons. *Law and Philosophy, 6*, 25–51.

Feinberg, J. (1965). The expressive function of punishment. *The Monist, 49*(3), 397–423.

Fletcher, L., & Weinstein, H. (2002). Violence and social repair: rethinking the contribution of justice to reconciliation. *Human Rights Quarterly, 24*, 573–639.

Gilbert, M. (1989). *Acting together*. London: Routledge.

Gilbert, M. (2001). Collective Remorse. In A. Jokic. (Ed.), *War crimes and collective wrongdoing*. Oxford Blackwell Publishing.

Gilbert, M. (2006). *A theory of political obligation: Membership, commitment, and the bonds of society*. Oxford: Oxford University Press.

Glasgow, J. (2015). The expressive theory of punishment. *Law and Philosophy*.

Govier, T. (1999). Forgiveness and the unforgivable. *American Philosophical Quarterly, 36*, 59–75.

Govier, T. (2002). *Forgiveness and revenge*. London: Routledge.

Greenspan, P. (1980). A case of mixed feelings: Ambivalence and the logic of emotion. In A. O. Rorty (Ed.), *Explaining emotions* (pp. 223–250). University of California Press.

Griswold, C. (2007). *Forgiveness: A philosophical exploration*. Cambridge: Cambridge University Press.

Hampton, J. (1992). Correcting harms versus righting wrongs: the goal of retribution. *UCLA Law Review, 39*, 1659–1702.

Hieronymi, P. (2001). Articulating an uncompromising forgiveness. *Philosophy and Phenomenological Research, 62*, 529–555.

Lippke, R. (2007). *Rethinking imprisonment*. Oxford: Oxford University Press.

List, C., & Pettit, P. (2010). *Group agency: The possibility, design and status of corporate agents*. Cambridge: Cambridge University Press.

May, L. (2004). *Crimes against humanity*. Cambridge: Cambridge University Press.

McNaughton, D., & Garrard, E. (2010). *Forgiveness*. Durham: Acumen.

Metz, T. (2000). Censure theory and intuitions about punishment. *Law and Philosophy, 19*, 491–512.

Metz, T. (2007). How to reconcile liberal politics with retributive punishment. *Oxford Journal of Legal Studies, 27*(4), 683–705.

Minow, M. (1998). *Between vengeance and forgiveness*. Boston: Beacon.

Moellendorf, D. (2007). Reconciliation as a political value. *Journal of Social Philosophy, 38*, 205–221.

Morris, H. (1968). Persons and punishment. *The Monist, 52*, 475–501.

Murphy, J. (2003). *Getting even*. Cambridge: Cambridge University Press.

Murphy, C. (2010). *A moral theory of political reconciliation*. Cambridge: Cambridge University Press.

Narayan, U. (1993). Appropriate responses and preventive benefits: justifying censure and hard treatment in legal punishment. *Oxford Journal of Legal Studies, 13*, 166–182.

Primoratz, I. (1989). Punishment as language. *Philosophy, 64*(248), 187–205.

Radzik, L. (2009). *Making amends: Atonment in morality, law, and politics*. Cambridge: Cambridge University Press.

Shapiro, S. (2011). *Legality*. Harvard University Press.

Stilz, A. (2009). *Liberal loyalty: Freedom, obligation and the state*. Cambridge: Cambridge University Press.

Tadros, V. (2005). *Criminal responsibility*. Oxford: Oxford University Press.

Vernon, R. (2002). What is a crime against humanity? *Journal of Political Philosophy*, 231–249.

Walker, M. U. (2006). *Moral repair*. Cambridge: Cambridge University Press.

Wringe, B. (2006). Why punish war crimes? Victor's justice and expressive justifications of punishment. *Law and Philosophy, 25*(2), 159–191.

Wringe, B. (2010) War crimes and expressive theories of punishment: communication or denunciation? *Res Publica, 16*(2), 119–133.

Wringe, B. (2012). Collective agents and communicative theories of punishment. *Journal of Social Philosophy, 43*, 436–456.

Wringe, B. (2013). Must punishment be intended to cause suffering? *Ethical Theory and Moral Practice, 16*, 863–877.
Wringe, B. (2015). Perp walks as punishment. *Ethical Theory and Moral Practice, 18*(3), 615–629.
Wringe, B. (2016a). *A Communicative Theory of Punishment*. London: Palgrave.
Wringe, B. (2016b). Rethinking Expressive Theories of Punishment: Why Denunciation is A Better Bet Than Communication or Pure Expression. *Philosophical Studies*