

# Journal of Practical Ethics

❧ Volume 5, Number 1. June 2017 ❧

# CONTENTS



Structural Injustice and the Place of Attachment <i>Lea Ypi</i>	I
Accepting Collective Responsibility for the Future <i>Stephen M. Gardiner</i>	22
The Death Penalty Debate: Four Problems and New Philosophical Perspectives <i>Masaki Ichinose</i>	53

*Editors in Chief:*

Roger Crisp (University of Oxford)  
Julian Savulescu (University of Oxford)

*Managing Editor:*

Dominic Wilkinson (University of Oxford)

*Associate Editors:*

Tom Douglas (University of Oxford)  
Kei Hiruta (University of Oxford)  
Guy Kahane (University of Oxford)  
Hannah Maslen (University of Oxford)

*Editorial Advisory Board:*

John Broome, Allen Buchanan, Tony Coady, Ryuichi Ida, Frances Kamm,  
Philip Pettit

*Editorial Assistant:*

Miriam Wood

The Journal of Practical Ethics is available online, free of charge, at:  
<http://jpe.ox.ac.uk>

*Editorial Policy*

The *Journal of Practical Ethics* is an invitation only journal. Papers are anonymously appraised prior to publication by expert reviewers who are not part of the editorial staff. It is entirely open access online, and print copies may be ordered at cost price via a print-on-demand service. Authors and reviewers are offered an honorarium for accepted articles. The journal aims to bring the best in academic moral and political philosophy, applied to practical matters, to a broader student or interested public audience. It seeks to promote informed, rational debate, and is not tied to any one particular viewpoint. The journal will present a range of views and conclusions within the analytic philosophy tradition. It is funded through the generous support of the *Uehiro Foundation in Ethics and Education*.

*Copyright*

The material in this journal is distributed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported licence. The full text of the licence is available at:

<http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>

© University of Oxford 2013 except as otherwise explicitly specified.  
ISSN: 2051-655X



# Structural Injustice and the Place of Attachment

LEA YPI

*London School of Economics and Political Science*

## ABSTRACT

Reflection on the historical injustice suffered by many formerly colonized groups has left us with a peculiar account of their claims to material objects. One important upshot of that account, relevant to present day justice, is that many people seem to think that members of indigenous groups have special claims to the use of particular external objects by virtue of their attachment to them. In the first part of this paper I argue against that attachment-based claim. In the second part I suggest that, to provide a normatively defensible account of why sometimes agents who are attached to certain external objects might also have special claims over them, the most important consideration is whether the agents making such claims suffer from structural injustice in the present. In the third part I try to explain why structural injustice matters, in what way attachment-based claims relate to it and when they count.



## I. INTRODUCTION

In early August 2013, an Australian mining company was fined \$150,000 for desecrating and damaging “Two Women Sitting Down”, an allegedly sacred site in the custody of the aboriginal people of Kunapa living near Tennant Creek, in the Australian Northern Territory. The damage was ascribed to blasting at a nearby mine that caused the collapse of an overhanging rock and the consequent split of the site into two parts, provoking irremediable damage to it. Although aboriginal representatives had been consulted when undertaking the works, the mining company was

accused of “abusing their trust” and concealing the full extent of the impact of the operation. The collapse, according to the Kunapa people, made it impossible for them to continue perceiving the site as part of a traditional songline (in this case narrating the story of blood spilled during the fight between a marsupial rat and a bandicoot) making it “much harder for Aboriginal people to recognise the dreaming” (see Jabour, 2013).

Many people tend to think, and I agree, that the verdict of the Australian courts in this case was justified. The Kunapa people had been affected in some non-negligible way; indeed, they had been wronged. Many people also think that the Kunapa were wronged *because* people who are attached to particular external objects ought to have a special say on how those objects are used. Both the diagnosis of the wrong and its suggested remedy, I want to argue, is misguided. Although attachment to particular external objects is a very important component of our descriptive explanation of why people object to the way others (who may not share the same sense of attachment) make use of such objects for other purposes (including distributive purposes), it does not help us construct a plausible account of *why* that complaint is normatively defensible. To provide a normatively defensible account of why sometimes agents who are attached to certain objects might be granted special claims over them, a more promising route is to ask whether agents making such claims suffer from structural injustice in the present. The first part of this paper explains and defends that claim. The second part illustrates the implications of my position. It suggests that while attachment is irrelevant for grounding special claims, it might have a role to play in determining the content of remedial obligations. The third part examines some objections.

## II. CLAIMS BASED ON ATTACHMENT

One reason for why attachment-based claims appear promising is that they seem to give us reason for understanding why *particular* agents might have special claims over *particular* external objects – an issue that seems difficult to settle by endorsing a conventionalist account that makes the allocation of objects to people depend on institutional norms that may or may not reflect the *particularity* of the relation.<sup>1</sup> The Kunapa people are connected to the rocks of Tennant Creek in some unique way:

1. For some recent discussions of the problem of particularity as related to attachment see Armstrong (2014), Stilz (2014), Moore (2015, , ch. 3 ff).

such objects are central to make sense of who they are, to the pursuit of purposive activities together with others they are related to, and to the distinctive system of rules and norms that they have reason to value. Thus, external objects are thought to matter because to be attached to those objects the way, say, Kunapa people are, implies to structure an entire life around activities sustained by access to those objects, and to do so in a way that recognizes and supports the meaning and values of certain group practices. To deny the Kunapa a special access to such objects central to their life plans, implies interfering with their autonomy to construct these lives as they see fit. This understanding of the role of attachment is implicit in the multicultural demand for the recognition of group-differentiated rights required to protect cultures as the context of choice in which individual autonomy often takes shape (see for example Kymlicka, 1989; Appiah, 2005; Raz 1994, esp. ch. 8).

Cultural attachment takes us to the problem of whether special claims to particular external objects can be justified with reference to the role that such objects play in promoting particular cultures and in sustaining the life plans of individuals whose lives are shaped by this belonging. Claims based on cultural attachment are often defended as pro-tanto claims, claims that are typically granted *some* validity but that can also be overridden. But under what conditions can they be overridden and when is it more difficult to do so? The crucial question, I want to argue, is whether agents making special claims to certain objects on grounds of cultural attachment suffer from structural injustice in the present. To introduce my argument, let me begin by contrasting two examples.

(i) *Seal-hunting*

The first example is the prohibition on seal-hunting. A directive from the European Council bans the import of products drawn from the hunting of seal pup, including luxury clothing, bags and jewellery (see European Commission, 1989). The ban, which affects a number of non-EU countries involved in the commerce of seal-skin, including Russia and Canada, is fiercely opposed by a number of fur traders but explicitly exempts products obtained by traditional indigenous communities such as the Canadian Inuit. Animal rights activists were pivotal to the enforcement of the ban and engaged in sharp debates rejecting the claims of fur traders. However, they also supported the exemption in the case of the Canadian Inuit. As the executive director for the Humane Society International Canada, an association which campaigned in favour of the prohibition of commercial seal hunt, put it: “we have

always argued that there should be an exemption for products from traditional Inuit hunters” (CBC News, 2014).<sup>2</sup>

It is tempting to argue that the case in favour of the exemption can be grounded on an argument from attachment: some traditional practices (e.g. seal-hunting) are central to the way particular indigenous groups structure their activities, promote particular ends and sustain a sense of who they are. Indeed, this is exactly how the animal-rights activists who supported the campaign in favour of the ban supported the exemption: they explained that they recognised that the tradition of seal-hunting played a unique role in the life of Inuit groups and that it would not be justifiable to deprive such groups from access to resources they were attached to and that were crucial to sustain traditional practices. This sounds plausible at first sight. But now contrast seal-hunting with a second example: fox-hunting.

(2) *Fox-hunting*

Fox-hunting is an activity, which involves the chasing, hunting and killing of foxes with the use of scent-hounds. It was practiced in England for some 300 years before coming to an end with the introduction of the Hunting Act in 2004, legislated by the then Labour government. The ban, which the House of Lords refused to approve despite an overwhelming majority voting in favour of it in the House of Commons, followed a controversial campaign against the ban, which listed among its supporters several celebrities, the then Conservative Party Leader Iain Duncan Smith and even members of the royal family (Anderson, 2006, p. 727). Pro-hunting activists insisted during their campaign on the importance of fox-hunting to sustain traditional English values and a rural way of life, increasingly threatened by city elites. Indeed, the vision of the Countryside Alliance, the pressure group established to mobilize against the hunting ban, was to promote “equal access to those facilities enjoyed in urban communities and where people can pursue their business, sports and pastimes according to the dictates of their own conscience, and in a society that appreciates and understands their way of life” (Countryside Alliance, 2013 [accessed]). Notwithstanding these protests, the ban was supported by many Britons who saw fox-hunting as a social practice essential only to sustain an aristocratic pastime which reproduced divisive class distinctions (see Burns et al, 2000). Oscar Wilde’s reference to “the English country gentleman galloping after a fox” as “the unspeakable in full

2. The controversy is still ongoing due to difficulties with the implementation of the exemption and the trouble with separating the Inuit harvest from that of East Coast seal hunters.



pursuit of the uneatable” seemed to capture well the mood of the public over the character of those who were interested in fox-hunting and their reasons for it (Wilde, 1998, p. 106). It was therefore no surprise when, several years after the legislation was passed, and despite all the lobbying efforts designed to change the legislation, surveys showed that 71% of the British public continued to express support for the ban (Wainwright, 2012). Despite all that, the Conservative Government led by Prime Minister Theresa May pledged on the wake of the June 2017 snap general election to give MPs a free vote on the repeal of the Hunting Act (see Bekhechi, 2017).

Fox-hunting and seal-hunting are in many ways similar. Members of groups attached to such practices have preferences over the use of external objects that are challenged by others. There is a clear tension concerning whose claims to prioritise. In both cases, an argument about attachment seems to be in play. In both cases, attachments have an important collective dimension that affects individual pursuits: they are central to a particular way of life, sustained by particular structured activities. In both cases such activities cannot be easily replaced by others without some loss to the customs and traditions of that particular social group. Interference in both of them involves some degree of restriction of agents’ autonomy. But why are we not as concerned about a bunch of aristocrats losing their privileges over activities essential to sustaining their life style as we are about indigenous people’s claim to land and objects central to their pursuits? If the answer were just an argument from attachment, it should have similar force in both cases. However, that argument needs to be further scrutinised.

Before proceeding with the main claim, it might be worth reflecting on one preliminary objection to the similarity between the two examples introduced above. One could argue that there is a difference between the projects and pursuits that are valuable to sustain a certain culture (like the culture of particular indigenous groups) and different ways of orienting one’s preferences in a society made up of different social groups (as with members of particular social classes). But it is too simplistic to dismiss the fact that there might be a distinct and pervasive culture associated to class as well as ethnic belonging. Think about all those 19th century novels (from Dickens, to Stendhal, from Tolstoy or Balzac) that narrate the difficulties that members of particular social classes encountered in trying to climb up the ladder of social hierarchy because of the fundamentally distinct habits, activities and practices associated to the practices of distinct social classes. Events like the French Revolution really did deprive the upper classes of objects (both land and resources) to which they were

significantly attached, and which played an essential role in sustaining their identities; to doubt this would be to profoundly misunderstand the impact of such events. Access to particular objects played a central role in the location of aristocratic life plans and depriving specific upper-class groups from having a special say on how resources were used significantly impaired their ability to make projects for the future thus interfering with their cultural context of choice. But how much should we worry about it?

Of course, some might worry about the process through which certain decisions were enforced, about the fact that in extreme cases they tended to involve the execution and imprisonment of members of aristocratic families. We might also worry that the suddenness with which certain events unfolded left many people hardly able to adjust to new circumstances - think about the scene in *Doctor Zhivago* where Yuri returns to Moscow on the eve of the Bolshevik revolution and is told by his wife Tonya that they now have to get used to live in a small space because many of the rooms in their big family house had to be offered to the new agricultural academy.<sup>3</sup> But all this has little to do with the fact that attachment to certain objects gives those who are related to such objects special claims on their use. So why are we more sympathetic to the claims of members of some groups but not others?

### III. INJUSTICE AND SPECIAL CLAIMS TO EXTERNAL OBJECTS

The claims of fox-hunters, like the claims of Queen Elizabeth to the parks, forests and animals surrounding Balmoral Castle seem importantly different from the claims of the Inuit and the claims of the Kunapa people over the rocks of Northern Australia. But if we ask ourselves why, we would struggle to find an adequate answer in an argument from cultural attachment and the centrality of particular objects to the pursuit of particular ends and activities. What might be an alternative account? I think we would be on stronger ground if we ceased to look for an argument that links particular external objects to the preferences of particular people and if we focused instead on the relations between people themselves. The case for paying *more* attention to the special claims of aboriginal groups and indigenous people than to Her

3. As one of the characters commenting on the events puts it: "This new thing, this marvel of history, this revelation, is exploded right into the very thick of daily life without the slightest consideration for its course. It doesn't start at the beginning, it starts in the middle, without any schedule, on the first weekday that comes along, while the traffic in the street is at its height." See Pasternak (1960, p.163), and in general all of chapter 6 for a good account of the disruption.

Majesty and her relatives does not rest on the special place that particular external objects occupy in sustaining their way of life, nor does it rest on the modality of relation to such objects or on their significance from the point of view of the claimants. From that point of view, there is no difference between aristocratic life plans and those of aboriginal communities. Both are the result of circumstances with which members of such groups strongly identify. Both play a significant role in how agents see themselves and their lives, who they regard as their peers and how they share with particular others specific values and purposive activities. And an abstract principle of equal respect commands concern for both. If we grant that in both cases the claim from attachment can be overridden, it should be overridden in both.

However, there are also important differences. While members of the first group have traditionally been implicated in systems of rules that they have coercively imposed on others, members of the second group have been victims of injustice and ongoing oppression whose effects persist in the present (Barry, 2001, ch. 7). Unlike privileged upper class representatives, indigenous people and members of aboriginal groups have found themselves on the receiving side of an endless chain of murder, persecution, abuse, exploitation and oppression with ongoing present-day implications for their relative power position in the societies of which they are members. It is in virtue of their subjection to injustice perpetuated by an objectionable system of rules and the social structures it replicates over time, a system that they did not contribute to making and that they still struggle to endorse, that sometimes we owe members of particular groups a special say on the use of external objects.<sup>4</sup>

To understand these points, consider first a simple, abstract case, which will become more complex in the next section. For years and years, Victor, the local bully, has persecuted and abused his neighbour Nora: prevented her from exiting her house, forced her to work for him, frightened her with constant threats, and always insulted her with unbearably offensive barbs. As a result, Nora is very badly-off, her dignity has been insulted and her self-respect severely impaired. At one point, and after several attempts, Nora successfully rebels and Victor has a change of heart: he realizes his mistake and, genuinely upset, decides to apologize, make appropriate amends and promises to never torment her again. Now suppose that from this point on Nora and Victor have to decide on how to use a shared allotment. In the past, Victor enjoyed a veto over how such allotment had to be used, he decided which vegetables to grow, what to do with them, and how to invest the income derived from

4. I emphasise “sometimes”, because this argument will be qualified in the next section.

their sale. But now Nora says she would rather plant flowers than grow vegetables in the allotment. Should Victor listen to her? Or should they take turns? I think, if Victor is not starving and can get vegetables from elsewhere, he ought to grant Nora her special request. He owes it to her in virtue of the effects of their tainted history of past interaction in the present, to make up for the injustice and abuse she has been suffering for all those years, and to allow her to develop her own interests and pursue her preferences unthreatened by the fear of her bullying neighbour. Victor, after all, had a chance to grow vegetables for as long as he wanted, and now it's Nora's turn to decide how to use the allotment.

But I don't think it matters to settle the case in favour of Nora rather than Victor to know that Nora is attached to flowers and Victor is attached to vegetables, or that the activity of flower-planting is crucial to Nora's life plans and to the development of her projects. I think we would grant Nora her special claim over the use of the allotment, even if she didn't have any plans at all, or if flowers didn't feature in them, or if we didn't know about any of her projects, or if she kept changing her mind on what to do with the allotment every day. The grounds on which the case in favour of Nora is made have very little to do with what contributes to a valuable pursuit of life projects, or with one's purposive goals and activities or with one's emotional investment in those activities or with what one is entitled to as a matter of abstract consideration.

But suppose Nora and Victor have now both died, and it is their children who are next door neighbours. What should they do with their shared allotment? Victor's grandchildren would like to go back to growing vegetables and Nora's family still like to plant flowers. If continuing to grant Nora's family a special claim over the use of the allotment is important to ensure that they feel respected by their neighbours, that they remain integrated in the area, that they continue to be free from stigmatization and disrespect, and that they are not bullied by the family of Victor ever again, then they should retain their special prerogatives. If none of this represents an ongoing threat and the descendants of Victor and Nora are in a roughly similar position, they might consider a different way of making these common decisions, one that gives both parties a more strictly equal say and where special claims over external objects no longer play a normatively relevant role.

#### IV. STRUCTURAL INJUSTICE AND THE ONGOING EFFECTS OF PAST OPPRESSION

The example with which I concluded the previous section simplifies matters enormously. The real victims of past injustice are hardly ever free from the burdens of the past. The tainted history of interaction with their bullying colonial masters continues to affect the way these groups relate to each other, the position in society of their members, whether they suffer from path-dependent disadvantages and whether they have a de facto equal say in matters of common concern. Yet the answer to whether they should have special claims over the use of external objects is similar to that in the simplified example above. To the extent that past abusive systems of rule continue to have a profound and pervasive effect on the lives of members of formerly oppressed groups, such groups should be granted a special say over the use of external objects and resources. This is so because despite any good faith effort at treating as equals victims of past injustice, the effects of that injustice are present and persistent even if the intentions of current members of former colonial societies are now different from those of their ancestors.

To better understand this point we should think about the effects on particular social groups of what I have referred to with the term structural injustice. A structure is commonly defined as a set of rules and resources recursively implicated in the reproduction of social systems in a way that both presupposes and creates certain patterned constraints on agents' positions and on the degree of social and political power that they control.<sup>5</sup> Structural injustice is often understood as the disempowerment of members of particular social groups resulting from the subjection to formal and informal rules that systematically thwart their access to resources, opportunities, offices and social positions normally available to other groups. This limited or unequal access may be the result of a causal history of subjection to particular social and political institutions or it may be the effect of unintended consequences leading to the consolidation of corrupt structural rules which are in turn upheld as a result of either complicity or indifference by others. Structural injustice then indicates the patterned constraints resulting from the cumulative effects of membership in partic-

5. For similar analysis of structure see Giddens (1984, 2). Young (2001) criticizes Giddens for placing too much emphasis on agency and the intentional role of agents in changing social structures. For further discussion of the various components of this definition see Sewell 1992. For the emphasis on the question of power in explaining the distinctive wrong of structural injustice, see Forst (2014, ch. 1).

ular groups marked by persistent disadvantage in agents' ability to create and uphold the rules of a social system responsible for the relative degree of social and political power they have within that system.<sup>6</sup>

If this analysis is correct, members of indigenous communities might be victims of structural injustice resulting from the persistent effects of former colonial structures on the opportunities and social roles available to them, without any member of the current society intending for this to happen. Empirically this is not far from true. To take only two relevant examples, Aboriginal Australians and Native Americans typically have lower incomes, higher infant mortality rates, lower levels of education and less access to desirable professions than members of any other groups in their societies, which suggests the presence of a pattern of structural injustice with effects that are difficult to remove in the present. It is in virtue of this that their special claims on the use of particular external objects are normatively appealing.

It is important to clarify that although structural injustice may include forms of distributive injustice, it is not reducible to them. When we speak of distributive injustice we typically refer to instances of problematic inequality between the amount of goods, resources or opportunities<sup>7</sup> available to different individuals, given a counterfactually just distributive background. In condemning that inequality and in seeking to remedy it, we do not necessarily reflect on the distinctive features of the process through which the goods in question are produced and might even be indifferent to how the relations between the individuals entitled to different bundles of goods have come to be what they are. We might ask, of course, whether individuals behave responsibly or irresponsibly in claiming particular shares or we might wonder what distributive principle would govern the rules of the social structure if individuals could choose without knowledge about their particular social position. But in all these cases, the political problem of unjust relations is seldom treated as a fundamentally distinctive kind of wrong from, say, an accident of fate with distributive implications for the life of different groups of people.<sup>8</sup>

When we focus on structural injustice, we focus on something different and more basic than the allocation of goods following ideally just distributive principles. Structural injustice is not reducible to any long-term departure from such an ideally

6. My account here is indebted to Young (2001, pp. 14-15). See also the discussion of the distinction between agential and structural injustice in Haslanger (2012), esp. chs. 7 and 11.

7. Here I use these terms without committing to a particular metric of distribution.

8. For a similar critique of the distributive paradigm see Young (1990) and Forst (2014, ch. 1).



just distribution, it has to do with the processes and relations through which that distribution comes about. The different focus on the relational wrongs responsible for the emergence of unjust structural rules has important implications for how we think about remedying structural injustice, the status and epistemic position of victims, the relation between groups that have been historically privileged or disadvantaged, and the ascription of duties required to take remedial action. The solution here is not identifying how much an agent has compared to another against an ideal baseline of who is entitled to what. Priority is given to the construction of appropriate political relations able to prevent fundamental social structures becoming vehicles of ongoing oppression. This is not to say that a just distributions of claims to external objects is of no importance. But it matters derivatively, as part of the establishment of ownership regimes able to remedy the failures of previously oppressive arrangements and in virtue of their effects to the distribution of current power relations. Special claims to resources and the related advantages conferred to particular groups are important in so far as they contribute to restoring the social positions of these groups such that they are no longer at risk of oppression in the present.

From this point of view then, certain special claims to external objects are more difficult to override not because those who make such claims are attached to the objects and the practices related to them. They have force as part of a process of empowerment (whether political, economic or symbolic) that helps members of structurally disadvantaged groups overcome the effects of past oppression. That they happen to be claims about objects to which members of these groups are attached, as in the case of the Kunapa or the Inuit, might affect the nature of plausible remedies and the content of the duties corresponding to them but not the grounds on which special claims are made. Empowerment requires taking measures designed to ensure that victims of structural injustice are able to participate in a society of equals without risking the ongoing silencing, marginalisation or stigmatisation of their claims by other more powerful groups (see Young, 1989). Of course, in some cases, where there is a basic recognition of the authority of existing institutions, members of formerly oppressed groups might be able to articulate their views and obtain fair resources and opportunities within functioning political structures without need to attend to special claims. But in many other cases, the effects of the past unilateral imposition of different political structures, and alien social norms and conventions might be so profound that the recognition of special claims becomes part of a process of empowerment that is instrumentally or symbolically important to avoid the ongoing

non-consensual political incorporation of these groups.<sup>9</sup> In both cases, as I said, attachment might be descriptively important to see why members of particular groups demand access to this or that particular object and to establish the content of specific duties able to remedy structural wrongs. But the mere fact of attachment neither grounds such claims nor explains why they seem difficult to override.

To see this last point, consider the following example. Suppose that instead of asking to have a special say on the particular rocks of the Australian National Territory to which they are attached, the Kunapa people asked for special quotas for their children to access schools they have never been admitted to before. If accommodating that claim could help remedy the effects of historical injustice on their present-day condition as part of a process of conferring economic, political, cultural or symbolic powers of resistance to the structural injustice that endangers such groups, special claims are plausibly made on grounds of forward-looking aspirations rather than backward-looking attachments, a point to which I shall return .

But what are the limits of these claims? I emphasised that we should be prepared to consider special claims, *if* doing so helped to remedy particular groups' condition of structural injustice. But it is also important not to fuel further injustices towards members of other groups. If the claims of the Kunapa can be accommodated without modifying the structural relations of other similarly positioned citizens, then we should be sympathetic to them. But such sympathy extends to claims with a past-oriented character as much as a future-oriented ones . This suggests that whatever objects are claimed by victims of past injustice it is not the historical attachment to such objects that gives special reasons for *why* members of these oppressed groups have a unique claim to them. What grounds the special claims are obligations to ensure that the effects of such injustice are not persistent in the present. It is hard to say what these obligations entail exactly in any given case. Sometimes (very often) what is needed is access to the economic resources necessary to fight poverty, inequality and social marginalisation and special claims to external objects help groups preserve traditional economic activities (fishing, hunting etc.) that are also important in the present. Sometimes it is a voice (and where necessary a group differentiated voice and even extensive rights of self-government) that would restore the equal standing of members of previously oppressed group in political institutions and structures largely set up by past colonial masters. Sometimes it is the symbolic

9. For a discussion of the relation between non-consensual political association and domination, see Ypi (2013).



recognition of particular claims demanded to restore the sense of dignity and self-respect that would empower these groups against the threat of ongoing injustice.<sup>10</sup> None of these decisions can be made in the abstract and without full involvement of members of these groups themselves. But notice that not much in this process relies on knowing anything about the previous structure of entitlements (who is attached to what, who deserves what or who got there first). Acknowledging the obligation to remedy present injustice is owed less to the loss or damage or mishandling of certain objects to which some people are attached and more to the wrong of treating as equals members of particular groups (for example by failing to apply the same conventions about property that were applied to members of other groups) *and* for the resulting structural injustice that is fuelled by that failure. But attachment seems to play very little role in grounding the claims. Claims based on attachment are no stronger than alternative (past-oriented) justifications that connect particular agents to particular objects (such as desert, improvement or labour).<sup>11</sup> If attachment matters at all, it is only in so far as claims based on attachment help remedy the condition of structural disadvantage that victims of past injustice suffer in the present and enables us to further specify the particular method of correction of current unjust relations.

## V. OBJECTIONS

We might wonder here about the exact link between remedying structural injustice and obligations that requires us to grant members of certain groups special claims over the use of particular external objects. If attachment plays no role in linking particular agents to particular objects does it mean that members of formerly colonised groups can claim pretty much anything regardless of the grounds of their particular relation to what they claim? I emphasised that they can, *if* granting such claims helps overcome their position of structural injustice and that satisfying their demands is possible without causing more injustice than we are trying to overcome. This means that the particularity issue that theories of attachment (and more generally historical theories of justice) seek to solve is, from my point of view, unsolvable. How particular shares are linked to particular objects is largely a matter of convention and I doubt that it is possible to find an account of how agents came to acquire a special title over

10. See Forst (2013, ch. 12.) for a good account of these different types of claims.

11. I also think (though I do not argue the case here) that they are no stronger than potential (future-oriented) justifications based on our imagining a special connection to objects that we never accessed in the past (objects that play a role in our dreams or aspirations, for example).

certain shares that can claim superiority over any other. But while it might be both unproductive and unnecessary to reflect on what justifies such conventions at their source, it is imperative to think about how they affect justice in the world we have, in what way such conventions can be improved and whether outstanding injustices can be overcome.

This might seem to lead to a second problem: if granting special claims to the use of particular external objects is conditional on overcoming structural injustice, should one rather not welcome other equally (or even more) effective ways of obtaining the same goal that do not depend on what indigenous people want? What if instead of granting the Kunapa a special say on the rocks of the Northern Territory or conceding the Inuit the right to hunt seals, they were provided with better opportunities for healthcare, education, and overall more resources enabling them to reacquire the political, economic and symbolic power necessary to overcome their condition? Wouldn't that be preferable?

I am in principle open to this suggestion but it comes with a few troubling features that are worth highlighting. A key problem here is what factors enable structural injustice to emerge and persist, who decides what counts as an effective remedy to it, through which processes that verdict is reached and who has a say in these processes. This is also why structural injustice speaks to a deeper and more basic problem than distributive injustice, a problem that is not indifferent to the reciprocal position of agents whose shares are affected. Imagine, to go back to my Nora and Victor case, that now Nora is back to planting flowers but Victor who (recall has had a change of heart) suggests that it would be much better for her if she joined in his activity of cultivating vegetables, perhaps taking some extra free courses on how to do it properly and having a greater share of the vegetables sold. And suppose that this is furthermore true. Since Victor has been the local bully for a while, he has managed to shape the preferences of other people in accordance with his. As a result, flower planting is now neither economically rewarding nor a particularly valued form of activity; it might just make Nora feel good but it does not help at all with improving her position in the neighbourhood. Growing vegetables on the other hand and selling them would be much more likely to bring Nora on a par with the rest of her neighbours and, in the long run, make her kids better off and more similar to everyone else. Now, I think we would still regard Victor's intervention just as a subtler instance of the same kind of dominating behaviour he has displayed all along. We would start doubting whether Victor ever really regretted what he did since what he ends up doing, yet again, is

telling Nora what to do and what is best for her, imposing values, standards and preferences on her which she does not really recognise as her own. And we would think so *even if* he does this with (what he thinks are) Nora's best interests in mind, and *even if* his strategy looks more successful in the long run. Even if we concede that there might be preferable alternatives other than following what members of previously oppressed groups tell us we should do in order to enable them to overcome structural injustice, it should trouble us that what those alternatives often imply ends up consolidating the same values, preferences and indicators that are at the basis of the system of rules that we are trying to improve or overcome. It is for this reason, I think, that we need to hear what victims of injustice have to say themselves about what they want (including where what they want is special access to particular objects). It is for this reason also that recognition of those special claims is often thought to play a key role in the instrumental and symbolic empowerment of such groups against the ongoing threat of structural injustice.

Finally, let me add a few words on the place of historical injustice in my argument. I have emphasised that the reason we care so much about historical injustice has to do with its effects on the lives of current generations of members of historically oppressed groups. The French committed grave injustices against the Spanish during the War of the Spanish Succession but we have no reason to be concerned with such injustices now (except for learning from and about them). History matters because events of the past have contributed to shaping the system of rules and resources that we have inherited and which is responsible for the generation of patterned constraints over the resources, opportunities and social positions available in the present. This is also where my argument differs from standard treatments of historical injustice, including those cases where authors are prepared to concede that historical injustice can be "superseded" (see Waldron, 1992). My strategy is not to acknowledge the existence of past-oriented claims to particular external objects or resources and then concede that they can be overridden by present-oriented ones.<sup>12</sup> My argument leads to reflect on the very category of *historical* injustice, questioning its normative significance when taken in isolation from current manifestations of

12. For a longer discussion, also of cases where wrongs of a territorial nature are at stake, see Ypi (2014).

wrongful structural relations. This leads to more plausible guidance on the grounds on which claims to supersession can be made.<sup>13</sup>

But here one might ask whether there is any principled difference between members of groups that have been exploited and oppressed throughout history and more recent victims of injustice, say members of particular immigrant groups facing structural threats of domination, discrimination and stigmatisation. If we are prepared to concede special claims over the use of external objects to indigenous people, conditional on such claims helping us remedy structural injustice, should we also be prepared to recognise special claims to other groups who are threatened in the present? If instead of the Kunapa people a group of recently arrived Somali immigrants laid a claim to the rocks of the Northern Territory (say because they always aspired and dreamt about living there) would we be prepared to endorse their request? I cannot see any strong reason to deny their claims. If the injustice from which these groups suffer is equally grave and pervasive, and if providing access to particular external objects helped remedy those injustices then we could not grant to one particular group privileges that are denied to the other. But what if there was a conflict between two equally oppressed group (say the Kunapa people and the Somali immigrants) over who can make special claims to the use of the Northern rocks? Again here, the answer cannot be given simply with the help of an attachment theory. It depends on the availability of other means to overcome injustice, on the process through which these alternatives are formulated and on the participation to that process of members of oppressed groups and the balance of reasons given for any set of preferences. There is very little we can say in advance of that process about who has more or less claim to what. What we can do instead is consider the relative position of those who advance such claims and the extent to which structural injustice affects their lives in an ongoing way.

Therefore, there is no reason to single out attachment as the most important argument for granting special claims. What matters, as I argued above, is not the relation between people and certain objects but the relationship of people to each other. What grounds the special claims to the use of external objects is the fact that agents making such claims suffer from structural injustices that might be remedied if such agents are empowered (both from a material and from a symbolic perspective). Structural injustice, as I argued, stems from the replication of system of rules and re-

13. The absence of a link between historical injustice and ongoing structural oppression is why the claims of indigenous people are still relevant and cannot be easily superseded, a weakness in existing theories about supersession.

sources with pervasive effects on the opportunities, offices and social roles available to members of particular groups. If the recognition of special claims is required to remedy such injustices, then these claims would be justified. The same does not apply if other groups, who are not victims of structural injustice (think about my example of aristocrats above) were to make such claims. But if we single out attachment as the most important ground for recognising special claims, we have no principled way to resist that extension.

I began this paper by conceding that there might be a *pro-tanto* argument in favour of attachment-based claims and developed the argument by reflecting on the conditions under which such claim might be overridden. But some might object that in the course of developing my account, I ended up weakening the force of attachment-based claims so much that it is now unclear whether attachment-based claims really matter at all, whether they even have the *pro tanto* weight that I conceded they might have. I find this critique plausible. Developing it further would lead us to question whether attachment grounds an even *prima facie* claim to external objects. To go back to our initial Kunapa case: why should the Kunapas' past relation to the rocks matter more than, for example, the aspiration to live there of the recently arrived group of Somali immigrants? Do backward-looking attachments really matter more than forward-looking aspirations? If the deeper argument on which attachment-based claims rest is the value of respecting agents' autonomy in making life plans, this applies as much to objects related to one's past life-plans as to future ones. Notice however that although I find this more demanding critique of attachment plausible, it is not necessary to endorse it to share my core argument. The main claim of my paper is that attachment-based considerations are harder to defeat when structural injustices are in play than when they are not. A more radical critique of the *pro-tanto* weight of attachment-based claims is compatible with my argument (and I am in fact sympathetic to it) but it is not necessary to endorse the main claims I made.

Before reaching a conclusion, three more clarifications are necessary. Firstly, in saying that structural injustice is crucial to see how some special claims to external objects by particular agents might be linked to the content of the duties to remedy structural injustice, I only mean to explain why we might occasionally be sympathetic to cases like that of the Australian Kunapa with which I started (even when we deny the distinctive role of attachment). But by invoking that alternative account, I do not mean to suggest that we should turn to considerations of structural injustice as providing defeasible reasons that *always* count in favour of granting members of

particular groups special claims over external objects. Other factors also need to be taken into account, and while structural injustice is a very powerful consideration, one would need to know more about other complicating features before coming to an all-things-considered assessment. Therefore, it is possible to formulate my argument in a more conditional way: special claims to external objects might not *always* matter but *if* we believe they do, we will be on stronger ground by invoking duties to remedy structural injustice than by appealing to an attachment-based structure of entitlements.

One other remark is in order here. Arguments that bear on the question of whether groups have special claims over external objects must often confront the thorny question of the relation of individuals to the groups of which they are members. However, it seems plausible that the question of whether members of a particular group can make special claims over the use of external objects in virtue of their position in society is independent of other considerations pertaining to the dynamic of justice within the group (i.e. whether the group oppresses such members and whether justice norms are internally followed). These issues require an independent assessment and in no way interfere with the question of whether members of that group are entitled to special claims given their position of structural disadvantage; how one should balance these considerations with considerations of injustice within the group is a separate matter.

Finally, notice how I began this paper by arguing that attachment to specific objects seems normatively important to address the so-called *particularity* question, i.e. to explain how particular agents relate to particular objects (including particular land, particular territory and particular resources) and why it would be pro tanto wrong to deny them special claims over the use of such objects. I also stressed that, if my argument above is correct, there is no normatively plausible answer to the particularity question.<sup>14</sup> The reason certain agents end up with certain shares is just an accident of history determined (if we are lucky) by convention: there is no nobler story to be told. All we can do is revisit those conventions with the aim of remedying the injustices they inherit and preventing the creation of new ones. Attachment, as such, matters as little to our claim to particular distributive shares as the other ways of connecting particular agents to particular objects that have often been the object

<sup>14</sup>. Elsewhere I have examined this question with regard to the problem of colonialism and the occupation of particular territories, trying to explain that the wrong of colonialism does not consist in depriving particular agents of territorial entitlements but in the way through which particular justice-based norms with territorial implications are established and enforced. See Ypi (2013).



of stark critiques (e.g. claims based on labour, desert or improvement). What really matters to justify special claims is not the relation between people and things but the relation of people to each other and the constraints of justice that shape that relation.

## VI. CONCLUSION

Attachment-based claims to particular external objects have received a great deal of attention. In this paper I argued that, absent certain considerations of background justice, attachment to external objects is normatively questionable. I also argued that a more plausible reason for granting members of particular groups special claims over the use of external objects has to do with remedying the structural effects of past injustice on their present-day condition. Attachment is easily overridden in grounding special claims but it might play some role in orienting our decisions about the content of the duties required to remedy structural injustice. If granting special claims is necessary to empower oppressed groups and free them from the threat of ongoing structural injustice, such claims need to be taken into account. If they play no such role, attachment-based claims are on much weaker ground (and might not even matter at all).

*Acknowledgments: I am very grateful to Arash Abizadeh, Marcus Agnafors, Chris Armstrong, Cecile Fabre, Rainer Forst, Bob Goodin, Matthew Lindauer, Catherine Lu, Margaret Moore, Liam Murphy, Sam Scheffler, Cord Schmelzle, Annie Stilz, Laura Valentini, and Tim Waligore, for extremely helpful comments on previous versions of this paper. I would also like to thank the editors and anonymous reviewers of the journal for their crucial guidance with the final version.*

## REFERENCES

- Alison Anderson, "Spinning the Rural Agenda: The Countryside Alliance, Fox Hunting and Social Policy", *Social Policy & Administration*, 40: 722–738 (2006)
- Anthony Appiah, *The Ethics of Identity* (Princeton: Princeton University Press 2005)
- Chris Armstrong, "Justice and attachment to natural resources," *Journal of Political Philosophy*, 22 (1), pp. 48-65 (2014)
- Brian Barry, *Culture and Equality: An Egalitarian Critique of Multiculturalism* (Cambridge: Polity Press, 2001)

Mimi Bekhechi, "Voters are being turned off by Theresa May's fox hunting plans" *The Independent*, <http://www.independent.co.uk/voices/fox-hunting-election-2017-theresa-may-keep-the-ban-public-support-a7769871.html> (2017) (accessed 22 June 2017)

Burns et al, *Government Inquiry Commission into Hunting with Dogs*, para. 4.12, <http://web.archive.org/web/20090410213249/http://www.defra.gov.uk/rural/hunting/inquiry/mainsections/huntingreport.htm> (2000) (accessed 22 June 2017)

CBC News, "Canada, EU edge toward agreement on Inuit sealskin products", <http://www.cbc.ca/news/canada/north/canada-eu-edge-toward-agreement-on-inuit-sealskin-products-1.2798211> (2014) (accessed 22 June 2017)

Countryside Alliance, <http://www.countryside-alliance.org/ca/article/our-vision>, (accessed 10 September 2013) [no longer available]

European Commission, Directive 89/370/EEC, [http://ec.europa.eu/environment/biodiversity/animal\\_welfare/seals/index\\_en.htm](http://ec.europa.eu/environment/biodiversity/animal_welfare/seals/index_en.htm) (1989) (accessed 22 June 2017)

Rainer Forst, *Justification and Critique: Towards a Critical Theory of Politics* (Cambridge: Polity Press, 2014)

———, *Toleration in Conflict* (Cambridge: Cambridge University Press, 2013)

Anthony Giddens, *The Constitution of Society : Introduction of the Theory of Structuration* (Berkeley: University of California Press, 1984)

Sally Haslanger, *Resisting Reality* (Oxford: Oxford University Press, 2012)

Bridie Jabour, "Mining company fined \$150k for desecrating Aboriginal sacred site", *The Guardian*, <https://www.theguardian.com/world/2013/aug/02/mining-company-fined-aboriginal-site> (2013) (accessed 22 June 2017)

Will Kymlicka, *Liberalism, Community, and Culture* (Oxford: Oxford University Press, 1989)

Margaret Moore, *A Political Theory of Territory* (Oxford: Oxford University Press, 2015)

Boris Pasternak, *Doctor Zhivago*, New York: Pantheon (1960)

Joseph Raz, *Ethics in the Public Domain: Essays in the Morality of Law and Politics* (Oxford: Oxford University Press, 1994)

W. H. Sewell, "A Theory of Structure - Duality, Agency, and Transformation," *American Journal of Sociology* 98, no. 1 (1992).

Anna Stilz, "Occupancy rights and the wrong of removal," *Philosophy and Public Affairs*, 41 (4), pp. 324-356 (2014)

Jeremy Waldron, "Superseding Historic Injustice", *Ethics*, vol. 103 (1), pp. 4-28. (1992)

Martin Wainwright, "Most Britons support fox hunting ban, survey finds", *The Guardian*, <https://www.theguardian.com/uk/2012/dec/25/britons-fox-hunting-ban-survey> (2012) (accessed 22 June 2017)



Oscar Wilde, "A Woman of No Importance" in *The Importance of Being Earnest and Other Plays*, (Oxford: Oxford University Press), New edition (1998)

I. M. Young, "Equality of Whom? Social Groups and Judgments of Injustice," *Journal of Political Philosophy* 9, no. 1 (2001)

———, "Polity and Group Difference - a Critique of the Ideal of Universal Citizenship," *Ethics* 99, no. 2 (1989)

———, *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990)

Lea Ypi, "A Permissive Theory of Territorial Rights," *The European Journal of Philosophy* 22, no. 2 (2014)

———, "What's Wrong with Colonialism," *Philosophy & Public Affairs* 41, no. 2 (2013)

# Accepting Collective Responsibility for the Future

Stephen M. Gardiner

*University of Washington*

## ABSTRACT

Existing institutions do not seem well-designed to address paradigmatically global, intergenerational and ecological problems, such as climate change.<sup>1</sup> In particular, they tend to crowd out intergenerational concern, and thereby facilitate a “tyranny of the contemporary” in which successive generations exploit the future to their own advantage in morally indefensible ways (albeit perhaps unintentionally). Overcoming such a tyranny will require both accepting responsibility for the future and meeting the institutional gap. I propose that we approach the first in terms of a traditional “delegated responsibility” model of the transmission of individual responsibility to collectives, and the second with a call for a global constitutional convention focused on future generations. In this paper, I develop the delegated responsibility model by suggesting how it leads us to understand both past failures and prospective responsibility. I then briefly defend the call for a global constitutional convention.



1. Elsewhere, I analyze climate change a “perfect moral storm” that involves all these dimensions (and more). Here I focus on the intergenerational aspect. See Gardiner 2011a.

## I. WHERE ARE WE?

### DESIGN ISSUES

In my view, humanity faces a serious problem of institutional design. This claim may initially seem radical. Yet it has resonance in climate circles, including occasionally among leading figures in the political establishment. For instance, Mary Robinson, United Nations Secretary-General's Special Envoy on Climate Change, former United Nations Human Rights Commissioner and former President of Ireland, once wrote about climate change (Robinson 2008):

*“The scope of these problems – and of the action required to treat them – reach beyond previous human challenges. Yet in the sixteen years since the UNFCCC was signed, global negotiations have proceeded at a glacial pace. We have collectively failed to grasp the scale and urgency of the problem. Climate change shows up countless weaknesses in our current institutional architecture.”*

More bluntly, in the run up to the disastrous Copenhagen climate meeting in 2009, Connie Hedegaard, then Danish Minister for Climate and Energy, and subsequently EU commissioner on climate action, quoted in Von Bulow (2009), said:

*“If the whole world comes to Copenhagen and leaves without making the needed political agreement, then I think it's a failure that is not just about climate. Then it's the whole global democratic system not being able to deliver results in one of the defining challenges of our century. And that ... should not be a possibility.”*

In essence, the problem is that currently dominant institutions do not appear well-suited for addressing serious, paradigmatically global, intergenerational and ecological problems, such as climate change.<sup>2</sup>

How do we explain these shortcomings? In my view, one of the main reasons is structural. Conventional institutions tend to be dominated by short-term concerns,

2. Some maintain that the Paris Agreement changes this situation. However, there remain substantial reasons for concern (e.g., Oxfam et al. 2015; Milman 2015; Anderson and Peters 2016), some of which echo my criticisms of previous agreements (Gardiner 2004; 2011a).

to have an overly narrow focus on economic matters, and not to have been designed with global, intergenerational or ecological concerns in mind.<sup>3</sup> Prominent examples of such institutions include market systems and national governments with three-to-five year election cycles. Notably, in many contemporary societies these institutions have mutually reinforcing effects on each other, and also play major roles in shaping the incentives facing other institutions.

To some extent, the limitations of current institutions are predictable. First, there is a *historical argument*. Presumably, there is a strong tendency for the members of any generation to create, maintain and shape institutions that disproportionately deal with the concerns that face them, here and now, in the present. Consequently, it would not be surprising for the set of institutions that emerges over time to reflect this, and show a tendency to highlight perspectives or problems that substantially crowd out broader intergenerational concerns. Absent deliberate, sustained intervention this would be just what one might expect, historically-speaking.

Second, there is a *strategic argument*. There is a clear temptation for each generation to behave badly with respect to the future. In doing so, it can secure benefits for itself and avoid bearing costs, in ways that seem morally unjustified. Consider two clear threats: the tyranny of the contemporary, and the problem of moral corruption.

The first threat is a general ethical challenge that I call the *tyranny of the contemporary*. To see the worry more clearly, consider a simple (and highly idealized) model. First, imagine a sequence of nonoverlapping generations. Second, suppose that each generation must make decisions about goods that are *temporally dispersed*. One type —‘front-loaded goods’— are such that their benefits accrue to the generation that produces them, but their costs are substantially deferred and fall on later generations. Another type —‘back-loaded goods’—are such that their costs accrue to the generation that produces them, but their benefits are substantially deferred and arise to later generations. Third, assume that each generation has preferences that are exclusively *generation-relative* in scope: they concern things that happen within the timeframe of its own existence (Gardiner 2015).

In such a situation, each generation has some reason to engage in “buck passing”. It can secure benefits for itself by imposing costs on its successors, and avoid costs to itself by failing to benefit its successors. Moreover, absent other factors (such as

3. Note that I argue only that this is one of the main reasons. Other issues, such as the embedded power of fossil fuel interests and other defects of existing institutions in representing current people, also play a considerable role.

moral convictions) this reason may be decisive. Since any given generation is (by hypothesis) concerned only with what happens during its own time, it has a standing concern for intragenerational benefits and costs, but no obvious concern for intergenerational costs and benefits. Buck-passing is thus the default position.

Other things being equal, intergenerational buck-passing raises serious ethical questions. As a matter of substance, each generation is likely to oversupply front-loaded goods and undersupply back-loaded goods relative to reasonable ethical norms.<sup>4</sup> This is clearest in the case of the oversupply of front-loaded goods. It seems unethical for an earlier generation simply to foist costs on a later generation with no consideration for its interests: for instance, without any compensation and without its consent. However, it is also relevant to the undersupply of back-loaded goods. On the (modest) assumption that, other things being equal, any given current generation has an obligation to engage in at least some back-loaded projects, then each generation will fail in its duties to the future if it does not invest in such projects.

The challenge posed by intergenerational buck-passing is also serious from a relational point of view. Later generations are subject to the arbitrary and apparently unaccountable power of earlier generations. This raises basic ethical questions, including questions of fairness, exploitation, domination and political legitimacy.<sup>5</sup> For example, elsewhere I argue that we should worry that buck-passing sometimes amounts to intergenerational extortion (Gardiner and Weisbach 2016, ch. 4; Gardiner 2017a).

The substantive and relational worries are unsettling enough when considering buck-passing by a single generation. However, this is only one part of the problem. First, in a typical tyranny of the contemporary, there is *iteration*. The temptation of buck-passing arises for each subsequent generation as it faces decisions about what to do about temporally dispersed goods. Consequently, the threat is often played out over many generations. Second, this iteration threatens *accumulation*: severe and catastrophic outcomes become more likely as the effects of ongoing buck-passing accumulate in the futher future. Third, this may also lead to *escalation* of the moral problem. As future effects accumulate, they place pressure even on otherwise decent

4. In this paper, I do not defend a particular view on the content of these norms. My aim is to present the problem at a higher level of abstraction, in a way which can accommodate a wide variety of possible views about that content. I will, however, tend to presuppose that our responsibilities to future generations are significant, and that the past 25 years of international climate policy constitute a *prima facie* failure to meet those responsibilities.

5. See, e.g., Bertram 2009; Nolt 2010; Smith 2013. (N.b. I disagree with Smith's characterization of my intergenerational storm.)

later generations to engage in buck-passing behavior. In particular, a later generation faced with severe impacts may (correctly) feel licenced to impose otherwise unacceptable burdens on the further future through a right to self-defense. As a result, taken to extremes, a tyranny of the contemporary threatens a dangerous downward spiral in social conditions, material and ethical.

The second clear threat—the threat of moral corruption—is related. The temptation to pass the buck to the future puts pressure on the ways in which we think and talk about issues in the real world, encouraging us to accept framings that distort our perceptions of moral reality. Specifically, the tyranny of the contemporary threatens to infect our framing of problems, and the theoretical “solutions” that attract us, in ways which facilitate intergenerational buck-passing. Most obviously, we can misdiagnose issues in ways which obscure their intergenerational dimensions.

Elsewhere I have argued that this threat is manifest in the climate case. For instance, the traditional framing of the climate problem in international relations and economics involves nation states facing a conventional prisoner’s dilemma, or tragedy of the commons scenario. Though each government recognizes that it is in the collective interest of all countries to cooperate to restrict emissions rather than not, each also believes that it is in the individual interest of its own country to defect from the cooperative outcome. Since all think the same, they all defect, with the result that all suffer. On the traditional view, then, the problem is one of self-inflicted, self-destructive behavior.

I find this view overly optimistic. For one thing, the traditional framing typically *assumes* that *the government of each state reliably represents the interests of its own citizens in perpetuity*, or at least for many generations into the future. Often this assumption is implicit, and so goes unnoticed. Yet once identified it strikes me as deeply unrealistic. Indeed, arguably, it is so far-fetched that it does not even pass an initial “laugh test” of minimal plausibility. The traditional argument is thus “a wolf in sheep’s clothing”.<sup>6</sup>

For another thing, the traditional view struggles to account for the depth of the failure of past climate agreements, such as Kyoto and Copenhagen. In my view, it is more plausible to think that the defections of governments from serious climate action have nothing to do with the long-term intergenerational interests of their own people, and everything to do with the narrower concerns of the present. If this is correct, the failures of the past do not amount to “self-inflicted and self-destructive” behavior.

6. Note that I am not claiming that all would be well if the laughable assumption turned out to be true.

tive” behavior, but instead involve current actors inflicting risks of severe harm that fall disproportionately on future generations (including of their own people), other countries, and nonhuman nature. In disguising this, the traditional tragedy of the commons analysis is seriously deceptive, and the tyranny of the contemporary closer to the truth.<sup>7</sup>

### PROSPECTS

It is tempting to think that the tyranny of the contemporary is such a deep problem that it cannot be overcome: that our prospects are hopeless. Yet this conclusion would be too quick. One way to see this is to notice that there are multiple potential drivers of such a tyranny, and they have different implications (Gardiner 2014, pp. 302-303). Specifically, the assumption in the previous illustrative model was that “each generation has preferences that are exclusively *generation-relative* in scope: they concern things that happen within the timeframe of its own existence”. However, there are various ways to interpret this assumption, and also less stringent assumptions that would create less pure, but still recognizable, versions of a tyranny of the contemporary. Consider three prominent examples.

The first, most obvious assumption, is generational ruthlessness: each generation is concerned only with its own interests, narrowly construed, and is indifferent to all other concerns, including notably those of future generations and the rest of nature. Some theorists in international relations talk as if something like ruthlessness is true, at least at the level of nation states. Specifically, they think that states act only for the sake of their own national interest, fairly narrowly defined. As noted above, such theorists often (implausibly, to my mind) assume that states pursue their own national interest understood in a robustly intergenerational way.<sup>8</sup> Generational ruthlessness is a harsh assumption. Nevertheless, even it does not imply that all is hopeless. After all, many mainstream strategies to convince current people to act on global environmental problems attempt to do so by showing them that (contrary to their initial impressions) such action is in their (typically short-term and economic) interests. The thought is that “green is good, *for us* (as well as for them)” and this is

7. My diagnosis has proven to have considerable explanatory power over time. See Gardiner 2001; 2004a; 2006, 2011a; 2017b.

8. See also, Gardiner and Weisbach 2016, ch. 3.



sufficient to drive action. This thought is made plausible by facts such as the falling price of renewables, and the indirect costs of fossil fuel consumption (e.g., to health).

Still, such “win-win” approaches have a notable weakness: they are hostage to fortune in notable ways. Most obviously, if it turns out not to be true that “going green” is in the short-term economic interest of current people (and especially maximally in their interests compared to other, browner options), the strategy of accommodating ruthlessness has nothing to say. Less obviously, the bar here is set very high and so not only accommodates, but also encourages extortion (Gardiner 2017a).

A second, in some ways less harsh, diagnosis is that the current generation is driven by *shallowness*, and in particular a fragile and ultimately mistaken conception of self-interest that obscures better and deeper visions. (For instance, classic versions of this diagnosis emphasize an unhealthy emphasis on consumerism, fed by excessive and manipulative advertising.) According to the shallowness account, the tyranny of the contemporary is particularly tragic: shallow conventional values are genuinely self-destructive for us, as well as ruinous for others. If we would only embrace a better (deeper) set of values, we would again find ourselves in a “win-win” situation where green is good for all.

The shallowness diagnosis often fuels responses to environmental problems that advocate alternative “green” lifestyles that highlight the benefits to the current generation of pro-environmental behaviors (e.g., the health benefits of cycling and less meat-intensive diets). Moreover, the diagnosis encourages the (often implicit) optimistic thought that mere recognition of shallowness should constitute a large part of the cure. The current generation, the thought goes, should be highly motivated to seek solutions that enhance its well-being, understood in a suitably enlightened way.

Still, here too there are reasons to hesitate. For one thing, the shallowness approach threatens to presuppose very widespread, perhaps global, agreement on what people recognize to be good for them. Yet such agreement has, historically and philosophically, been hard to come by. For another thing, even if agreement can be found, recognition of our true interests is often insufficient to ensure adequate action. We already frequently fail to do what is best for us, and for a multitude of reasons. (For example, often we fail to eat well and exercise.) Moreover, sometimes the reasons for not doing what is good for us are themselves fairly shallow: for example, often we are simply stuck in our ways, due perhaps to laziness, or complacency, or just being uncomfortable with change.

In my view, both ruthlessness and shallowness contribute to the tyranny of



the contemporary in the case of climate change. There are therefore good reasons to continue to pursue the associated “win-win” strategies to at least some extent. Nevertheless, I believe that a third driver is even more important: *obliviousness*. Here the charge is that conventional institutions are blind, or perhaps hostile, to intergenerational concern.

On this diagnosis, current people may have genuine intergenerational values, including around the strong need to protect future generations from the negative effects of phenomena such as climate change. However, current institutions lack the wherewithal with which to *register* such concern and drive it towards an appropriate impact on policy. Instead, intergenerational concern is largely invisible to existing institutions, or else swamped by other concerns. Consequently, even though people may have genuine intergenerational concern, existing institutions make it appear *as if* the assumption of generation-relative motivation were true. In other words, there is a profound institutional gap between the plausible values of the population and the forces that ultimately drive policy. Unless this gap can be filled, the tyranny looks set to run and run.

Some will doubt that the current generation has genuine intergenerational concern, or perhaps sufficient levels of such concern. Though I am optimistic that they do, or can at least be roused so that they do, in the end this is an empirical matter. I will say that genuine intergenerational concern is needed if we are to overcome the *moral* problem posed by the tyranny of the contemporary in climate and other cases. In particular, even if the *practical* problem of climate change were resolved by showing that the interests of the current generation do in fact perfectly align with those of future generations (as the more mainstream responses to ruthlessness and shallowness attempt to do), such a resolution would demonstrate only that the current generation is capable of pursuing its own interests more effectively in situations that turn out *not* to be true instances of the tyranny of the contemporary.

As I said above, in my view, overcoming the tyranny of the contemporary in a case like climate change will require both accepting responsibility for the future, and meeting the institutional gap. I will now propose that we approach the first (accepting responsibility for the future) in terms of a traditional “delegated responsibility” model of the transmission of individual responsibility to collectives, and the second (meeting the institutional gap) with a call for a global constitutional convention focused on future generations.

## II. DELEGATED RESPONSIBILITY MODEL

Let us begin with a basic sketch of the delegated responsibility model (DRM).

### A BASIC SKETCH

Individual agents have moral and political responsibilities.<sup>9</sup> Many can be discharged by individual action. However, others are best discharged collectively, including a large number that can *only* be either fully, or successfully, or adequately discharged collectively. These include many paradigm cases of moral and political responsibility. In such cases, individuals have duties to cooperate with others to discharge their (individual) responsibilities collectively. Typically, this is done through a delegation of authority by the relevant individuals to a collectively-sanctioned institution. This then acts “in their name”. Notably, without delegation, the collective institution lacks authority, and the original responsibility remains with the individuals. Similarly, if a collectively-sanctioned institution is established but fails to discharge the relevant responsibilities, then responsibility returns to the individuals again, to either fix the institution, or to find some other way to discharge the responsibility (e.g., typically by setting up some new institution).

As I interpret it, the delegated responsibility model posits two levels of responsibility. Call these ‘level 1’ (L<sub>1</sub>) and ‘level 2’ (L<sub>2</sub>). We might think of them as the *origin level* and the *delegated (or discharged) level*. On the delegated responsibility model, level 1 (the origin level) is the individual level; level 2 (the delegated level) is the collective level. Hence, the idea of the DRM is that individual agents have L<sub>1</sub> responsibility for action that they then delegate to another collectively-sanctioned agent (or agents) at L<sub>2</sub>, on the grounds that the L<sub>1</sub> responsibilities are best discharged collectively.<sup>10</sup> In paradigm cases, this is because they can *only* be either fully, or successfully, or adequately, discharged collectively, though in practice it is often merely because they

9. For present purposes, I will not distinguish between these categories. Again, my hope is that my current argument is at a higher level of abstraction, and compatible with a wide variety of views about the relationship between these two categories. Moreover, while my own position suggests that they are tightly linked, the details do not seem critical to the general argument being offered here.

10. Such L<sub>2</sub> agents are typically collective agents, but may also be individuals (e.g., if the need for collective discharging rests on a need for joint authorization, rather than for a distinctively collective agent). In addition, such delegation need not imply (and usually does not) that individuals do not retain some responsibilities at the individual level, nor does it imply that there are no new responsibilities for individuals in light of the delegation. For example, individuals may have significant (residual and/or new) responsibilities as individuals to work to sustain L<sub>2</sub> institutions.

are most efficiently, or conveniently, discharged collectively (e.g., because doing so dramatically reduced transaction costs, because relevant institutions already exist for other purposes, etc.).<sup>11</sup> The key point, however, is that collectively-authorized (L2) agents are empowered (and often also created) with level 2 responsibilities to discharge the relevant set of level 1 responsibilities of individuals. Notably, the L2 agents take their authority from L1 agents, and so from individuals. Consequently, individuals remain responsible for their delegation of responsibility to L2. This is part of what it means to say that L2 agents act “in the name” of L1 agents.<sup>12</sup> Their authority is at the delegated level, and therefore must be understood in terms of the delegation from the individuals who have the initial authority at the origin level.

#### STATUS OF THE MODEL

Before exploring the model, allow me to make three clarifications about its status. The first clarification is that although, on my understanding, the delegated responsibility model is a traditional one in Anglo-American political philosophy, this is not to say that it is a secure model. There are many ways in which it may be challenged, some serious, for example by approaches such as Iris Marion Young’s (Young 2013). Ultimately, the model may yet prove philosophically unattractive, or practically unworkable. Nevertheless, the DRM does have considerable explanatory power, both theoretically and when it comes to understanding a number of actual political practices. In what follows, rather than directly defending the model, I will focus on its explanatory power, with an emphasis on what the model might help us understand about global, intergenerational and ecological problems such as climate change.

The second clarification is that much of what I shall be arguing about the DRM is theoretically fairly modest, and should be of interest to a wide range of theorists of many different stripes. In particular, much of the explanatory power of the DRM is not undermined by deeper philosophical questions about the foundations and wider relevance of the model itself. To illustrate this, let me make two points.

11. Sometimes the reason will be that although individual discharging would be the most effective approach, it conflicts with other priorities. For instance, individuals may find such action irksome (e.g., as in the complaint that “socialism would take too many evenings”).

12. My distinction differs from Simon Caney’s contrast between first- and second-order responsibilities for climate change, and addresses different issues (Caney 2014, pp. 10-13). For instance, Caney’s second-order responsibilities are (i) “responsibilities to ensure that agents comply with their first-order responsibilities”, and (ii) include responsibilities to allocate new first-order responsibilities to those who previously lacked them (Caney 2014, p. 13). By contrast, (my) level 2 responsibilities are not focused on ensuring compliance, nor do they concern allocating (my) level 1 responsibilities.

First, accepting the DRM need not imply endorsing an underlying monism about responsibility. The DRM explains how a form of moral and political responsibility works. It need not claim to cover all kinds of moral and political responsibility. For instance, one might accept the DRM only as one part of a more comprehensive, pluralistic account of responsibility.

Second, there is a clear reason to think that a more comprehensive account is needed. The DRM takes the relevant form of individual responsibility as primary, and the relevant collective responsibility as derivative. Still, there is a complication. The delegated responsibility model does not by itself determine from where the original individual responsibility comes. It is a view about the *relationship* between (a form of) individual responsibility and (a form of) collective responsibility, not about the nature of (the relevant) individual responsibility itself. Consequently, it is compatible with a variety of views about the nature of the relevant individual responsibility.

The second point can be illustrated with two salient examples. One class of views takes the relevant sense of individual responsibility to rest on a *prior individualistic account*. For instance, some will likely claim that individuals assume responsibility (in a prior individualistic way) by expressly joining a group that is essentially constituted by a set of norms with the structure of the DRM model (e.g., some voluntary association). On this kind of view, the individual responsibility that is primary within the DRM is (from a wider perspective) itself derivative from a prior individualistic model. So, though allowing for a derivative explanation, such views are “individualistic all the way down”.

Another class of views, however, will likely take the individual responsibility that is primary within the DRM to derive from *a prior form of collective responsibility*. So, for example, some will claim that one has individual responsibilities for ensuring basic security in one’s own nation because one is a member of that nation (e.g., because one was born into it). Here, the primary responsibility from the point of view of the DRM is individual responsibility, but that individual responsibility is itself predicated on (a) group membership, and (b) an understanding of what it is to be a member of such a group that makes the DRM model and its invocation of individual responsibility appropriate. On such views, the individual responsibility cited by the DRM is in one sense a “conduit” between different forms of collective responsibility.

These examples show that, in terms of the terminology of levels, some views of responsibility will marry the DRM with a further view about the origins of level 1 responsibility. Some of these will argue that their view justifies positing further levels

below L<sub>1</sub> to account for this. Usually, this will be through the signaling of a 'level zero' (L<sub>0</sub>), though perhaps some will want to go deeper. Similarly, other views may elaborate on the DRM by positing levels beyond level 2 (e.g., level 3, level 4). One reason that this more complex structure may emerge would be if it takes an initial delegation to a L<sub>2</sub> entity for it to be understood that further institutional structure is needed but this does not replace the need for the initial L<sub>2</sub> entity. For example, one might imagine the creation of a number of L<sub>2</sub> entities at the local level, where this leads to a recognition that a national level institution is also needed, and in a way which does not involve a replacement of the L<sub>2</sub> (local) entities, but rather a supplement at a new level 3.

My third and final clarification concerns my own background view. While most of the argument I will be giving relies only on a modest endorsement of the DRM, takes no position on more general or foundational issues, and so should be of interest to a wider range of theorists, it nevertheless seems wise, in the spirit of full disclosure, to signal that I am sympathetic to a rival position to those just mentioned, and one that takes the relevant individual responsibility to be primary in a stronger sense. Specifically, I take the relevant kind of individual responsibility to be *foundationally sited in individuals simply in virtue of their moral agency*. In other words, responsibility is not sited in individuals because (for example) they have *acquired* responsibility in other ways, such as through their choices, or being members of groups, or becoming entangled in various forms of other relationships, whether causal, moral or of some other kind. Instead, on the view I have in mind, agency itself has its price. A heavy part of this price is the assumption of agency responsibility, the sort of moral and political responsibility that one has simply in virtue of being a moral agent in the first place.

Though I will not defend this approach in this paper, I will make two remarks to give a sense of where my version of the agency view is coming from and what motivates it. The first remark is that the agency view as I understand suggests a general outlook on humanity's situation in dealing with a challenge like climate change. At the most fundamental level, we are seven billion (or so) individual agents who find ourselves together facing a common challenge on this planet that only we can address and in some sense must address together. On this outlook, our basic situation is one of "we're all in this together (as agents)".

To adopt this outlook is not to deny that, looked at conventionally, our existing agency is structured through countless historical and current social practices, and

these limit our actions in various complex ways. Nevertheless, in my view the basic situation remains. In particular, though the conventional structures are deeply socially embedded and in many ways resilient to change (to good and bad effect), they retain much of their force due to *our ongoing support* of existing practices, support that is both practical and theoretical. In other words, the continued status (and therefore effects) of existing structures is largely contingent on our endorsement of them, as individuals and collectively. Consequently, a key part of the challenge facing us is how to understand their strengths and weaknesses, and what is at stake in changing them if they prove to be inadequate.

The second remark is that the “we’re in this together (as agents)” view fits into a broader starting-point that elsewhere I call the *Perspective of Humanity* (Cf. Gardiner 2011b):

*“The basic position of human individuals is that of members of a recently-evolved species on a small planet in an otherwise inhospitable solar system, amidst a vast, and currently unreachable universe. At present, the only viable (e.g., adequate, accessible) home for the human species and for all (known) others is the planet on which we reside. However, humans have attained unprecedented power over the planet, and now have the ability to influence profoundly the basic physical and ecological systems that support life as we know it, and especially for the worse.”*

In my view, this claim captures fundamental facts about our situation that must be respected by moral and political theories. These facts pose questions about how we are to understand ourselves and how we are to act. Moreover, it is natural to think that such questions are fundamental to social and political philosophy, in the sense that our responses to them *frame and limit* the answers we may give to other pressing social and political questions, such as how we are to live our own lives, and what we owe to contemporaries of our own species.

Though I signal the view and the perspective, I will not pursue them in more detail here. Fortunately, I suspect that little of what I have to say in the rest of this paper depends on them, and that proponents of other views can still accept the DRM, and its explanatory value. Still, I signal these background positions to forestall certain kinds of questions – generally questions that ask why I am so strongly emphasizing the DRM, and not pursuing the topic in a way that proponents of the other



two classes of position may prefer or expect, since they believe they have “deeper” individualistic or collectivist accounts on which to draw.

### III. EXPLANATORY VALUE

Let us turn now to the explanatory value of the DRM. I will identify five areas where the DRM is illuminating: in explaining silence, unfamiliarity, inhibition, evasion and dystopia.

#### SILENCE

The first two areas (silence and unfamiliarity) involve a challenge that is often posed to responsibility in situations like climate change. In a general form, the challenge is memorably described by Dale Jamieson (1992, p. 149), who says:

*“Today we face the possibility that the global environment may be destroyed, yet no one will be responsible. This is a new problem.”*

Jamieson thinks that our ordinary conceptions of moral and political responsibility fail to find traction when it comes to global environmental problems. Commonsense morality and the theories that rely on it are *puzzlingly silent* on these important issues.

There are a number of reasons for Jamieson’s claim, and I discuss many of them extensively elsewhere (Gardiner 2011c; Jamieson 2013; Gardiner 2013). Here, however, I will focus on just one: the idea that commonsense morality faces deep difficulties when serious ethical problems are created by us just “getting on with our lives” (Jamieson 2013, p. 42), or actions that are “usual” (Sinnott-Armstrong 2005). Jamieson glosses this point by saying (2014, p. 148):

*“Most of the time we do not subject what people do to moral evaluation. This may be because we consider most of what people do to be “their business”, belonging to a private sphere that is beyond the reach of morality. Or it may be because we regard most of what people do to be permissible. ... Various moral theorists would like to dislodge this way of seeing things, but nevertheless this is more or less how most of us see things most of the time.”*

In my view, one appeal of the DRM is that it has considerable explanatory power when it comes to making sense of the worries expressed by Jamieson and others<sup>13</sup>, yet without succumbing to the general skepticism about responsibility that some of their remarks suggest.<sup>14</sup>

Let us begin with the general challenge of ‘no one is responsible’. The DRM gives rise to two natural explanations of the ongoing policy failures with respect to climate change, and other intergenerational issues (Gardiner 2011c). According to the first, the DRM implies that *level 2 (delegated) responsibility has been assigned to national governments*. Although the initial level 1 (origin level) responsibility rests with individuals, they have delegated authority to their national governments to deal with global environmental problems. However, national governments are failing to discharge their L2 responsibilities; hence, they are subject to (moral) criticism. Consequently, national governments must rise to the challenge, and individuals (because of their L1 responsibility) must urge them to do this. If this project fails (and arguably it has failed in the climate case, at least for a good part of the last 25 years), then the attempt to delegate responsibility has been unsuccessful. Thus, we remain at L1, where individuals have the responsibility to find another way to solve the problem, presumably through another attempt to establish effective institutions. We might add that both individuals and national governments share in the responsibility for such failures, particularly insofar as failure was reasonably predictable, and especially if it was a calculated failure (e.g., for purposes of wasting time) (Gardiner 2004; 2011a).

According to the second natural explanation, the DRM implies that *no one has been assigned level 2 (delegated) responsibility*. On this view, national governments were not designed to deal with global, intergenerational, and ecological problems such as climate change. Hence, the assignment of responsibility to them is unwarranted, and

13. My claim is that the DRM provides one illuminating explanation. I do not claim that it is the only possible explanation, nor do I try (here) to defend it against alternatives.

14. An anonymous reviewer objects (1) that Jamieson’s point is about moral responsibility, (2) that moral and political responsibility are “apples and oranges”, but (3) that my approach mixes moral and political responsibility to generate its explanatory power, and so (4) does not address Jamieson’s primary underlying concern, which the reviewer takes to be “precisely (a) that no single agent might be to blame for climate change and (b) that no agent might be clearly identifiable as morally responsible”. I agree with at least some versions of (1) and (3), but do not think this is a problem, since I am not convinced by robust versions of (2). (Also, Jamieson explicitly objects to political responsibility as well in some places, so (1) is not as pertinent as it may initially appear.) In addition, as I understand it, my approach does address relevant versions of (4). Although the reviewer may be right that it does not do so within Jamieson’s framework as such (or the reviewer’s), I am not sure why that is a persuasive objection, rather than (say) an advantage of my (allegedly rival) framework.



perhaps also unfair.<sup>15</sup> (This view is the one that most resonates with Jamieson’s claim that “the global environment might be destroyed, yet no one will be responsible”.) According to the DRM, then, responsibility remains solely at level 1 (the origin level). Having proceeded no further, individuals bear responsibility, since no attempt to delegate was made, or all such attempts were unsuccessful, resulting in no level 2 assignment of responsibility at all. Either way, the level 1 responsibility remains.

In summary, taken together these two explanations seriously blunt the initial force of the general challenge. There is no theoretical mystery here; let alone one that presents a dramatically new problem for moral and political philosophy.<sup>16</sup>

#### UNFAMILIARITY

Let us turn then to the more specific version of the ‘no one is responsible’ worry: that climate change and problems like it fail to register with commonsense morality because they involve our “going about our everyday lives”, or indulging in merely “usual” actions or activities, or are insulated from moral criticism by a background social convention of most things not being morally accessible. Let me make three points in response.

The first point is that this worry appears to miss the mark, since the delegated responsibility model supplies a natural explanation. This explanation takes the form of an error theory. If we understand “going about our everyday lives” and the “usual” within the proper political context supplied by the DRM, the worry rests on a confusion. Specifically, part of the point of the DRM is to assign delegated (level 2 and higher) responsibilities in ways that help to *define* “the usual”, including by partly determining the extent of our individual freedoms to “go about our everyday business” without undue burdens. In other words, the notions of “the usual” and “going about our everyday lives” to which the objection appeals are *not prior* to the DRM and its delegation of L2 (and higher) responsibilities, but are instead bound up with that

15. Though there is some truth to this claim, it should be noted that most political actors have acted as if the role did belong to them and they were competent to discharge it, rather than (for example) declaring to their constituencies that the topic was outside of their purview, or advocating for fundamentally new or different institutions. Given this, some responsibility remains (cf. Gardiner 2011c).

16. Does my account beg the question against Jamieson, since it presupposes that individuals have responsibility at L1? I think not. (At least, it does not beg the question any more than presupposing that there is no such responsibility.) As we shall see, most of the explanatory power comes from the DRM model as such, not from the simple assumption of individual responsibility. For instance, the DRM provides a useful *error theory* as to why concerns about a lack of individual responsibility might initially arise even if such responsibility is actually present. (See below.)

assignment. Consequently, appeals to those notions are dubious in situations where the attempt to delegate has failed, including when no attempt to assign L2 responsibilities has even been made.

The second point is that it is not clear that a more demanding account of individual responsibility is so implausible, either at the level of the usual or more foundationally. In fact, it seems that demanding ideals may already be there, lurking in background conceptions of how political responsibility is supposed to work.

Consider a more standard case than climate.<sup>17</sup> Suppose there is a breakdown in basic security in a neighboring city in one's own country. For example, suppose one lives in Sacramento and the security forces in San Francisco are wiped out in a terrible terrorist attack. Who would have the responsibility to deal with it? Presumably, in the first instance it is the city and state governments, and (failing that) the government of the United States. Why? On the delegated responsibility model, it is because they already have delegated authority to act "in our name" to ensure basic security. However, what if all these efforts to delegate fail? Would the rest of us be off the hook? On the delegated responsibility model, the obvious answer is 'no'. Primarily, each of us would have some responsibility to try to get existing institutions to live up to the responsibilities delegated to them, and (if this turns out to be hopeless) to establish new ones to replace them.<sup>18</sup> Secondly, we would also have a responsibility not to thwart good efforts to achieve these goals, but to cooperate with them. For instance, we should not try to benefit from the lawlessness by sending in looting parties, or making deals with potential looters on the black market.

Would each of us also have a personal responsibility to "get armed and go West" in order to police the streets of San Francisco ourselves? In practice, I doubt that it would come to that. Uncoordinated individual action would be a pretty poor way of addressing the real problem, and come at a very high cost. If we got to the point where average individuals had seriously to consider packing rifles and flak jackets, an awful lot would have to have gone wrong. Moreover, there would also have to be a good chance of making a meaningful difference, and the prospects for better solutions would have to be bleak.

Nevertheless, in principle, if all other efforts towards better solutions failed, and if we could plausibly make a real and positive difference, the idea of robust person-

17. These next two paragraphs are adapted from Gardiner 2013.

18. The wider burden may initially fall on Americans. However, under a number of circumstances the 'us' would extend to a wider global public.

al responsibilities is not so strange, especially as a stop-gap measure until a better response could be put in place. Indeed, assumptions of personal responsibility are often made in such circumstances. For example, when communities develop response plans for predictable disasters, such as earthquakes in high risk areas, they often assume that individuals must initially pick up the slack, at least until a more effective collective response can be made.

If this is correct, on a plausible interpretation of the delegated responsibility model, the reason that this kind of individual responsibility seems unfamiliar is not that there is a *conceptual* problem with individual responsibility, but that focusing on the individual's personal behavior seems the wrong way to tackle the problem actually being faced, or at least so far down the list of serious options that it is a poor focus for action. Instead, in the cases with which we are most familiar political responsibility seems much more central.

Still, none of this implies that a demanding notion of individual personal responsibility is not already implicitly present in our lives, embedded in background assumptions about the way our society is supposed to work, and so, in a sense, constituting one part of the "usual". On the contrary, though normally "invisible" in daily life, robust forms of personal responsibility do persist in many contemporary societies as an important part of their organizational structure. Common examples include: jury duty, the draft, mandatory national service, and so on. Consequently, such responsibility is in fact more visible than the initial objection acknowledges. Severe circumstances expose it more clearly to the light, but they do not create a "new" problem.

The third point is that the feeling of unfamiliarity is predictable, especially in cases where delegation is perceived to be highly effective in other domains of life (Gardiner 2011c). In particular, some who endorse the DRM will also hold wider conceptions of political philosophy that embed a familiar socio-political ideal. According to this ideal, effective institutions should discharge as many ethical responsibilities as possible for citizens, so that individuals do not have to worry about such responsibilities, but become maximally free to engage in their own pursuits (subject to external constraints set out by the system). However, on this ideal, success tends to breed elimination of responsibility in personal life. Consequently, the more effective a social system is (or is perceived to be) at discharging responsibilities in general, the more demanding any significant unmet responsibilities will seem. In other words, for those used to broad freedoms to pursue their own ends without worrying about

wider responsibilities, the emergence of a serious failure to discharge is likely to be deeply jarring: the issues will seem very unfamiliar and the nature of the responsibilities extreme. Yet this may say more about the past successes of the DRM than current or future failures.

We can conclude that the DRM seems highly effective in countering the initial force of the “no one is responsible” and “getting on with our lives” objections. Still, there is more to be said. The DRM has additional explanatory force when it comes to understanding other obstacles to action. The issue of L2 assignment of responsibilities raises several uncomfortable possibilities. In many circumstances, actors will be tempted to take advantage of the DRM in a variety of (often interconnected) ways. Sadly, this temptation appears especially salient in the case of climate and other global and intergenerational responsibilities, when the threat of moral corruption is high.

#### INHIBITING

Let us begin with the thought that a lack of level 2 responsibility assignment, or at least of effective assignment, will be highly convenient for some actors in some circumstances. Suppose, for instance, that some actors do not want to see some problem solved. (These actors can include both L1 agents and L2 agents established for other purposes, such as discharging other L1 duties.) The DRM suggests several ways of inhibiting or disrupting the process so that effective delegation does not occur. Consider just three likely strategies.

We might call the first strategy ‘internal structural sabotage’. Actors may manipulate specific L2 responsibility assignments in the relevant domain so that they are ineffective. For instance, they may facilitate agreements that are highly voluntary, with weak incentives, accountability and enforcement mechanisms. For example, in the climate case, consider the structure of the Kyoto Protocol and the initial form of the Paris Agreement (Gardiner 2004; 2011a).

The second strategy is ‘external structural sabotage’. Actors may arrange L2 responsibility assignments for *other* issues in such a way as to create powerful L2 agents in those (other) domains that oppose effective L2 assignments in the area in which they wish to prevent action. For instance, L2 actors from other domains (such as national governments or corporations) may prevent collective organization from forming in the relevant area (e.g., global environmental protection), they may ensure that any new L2 actors are so weak that they cannot perform their function (e.g., the

UNFCCC), or they may work to undermine those actors once they are in place (e.g., by establishing international trade agreements that prevent trade sanctions, and so inhibit strong environmental regulation).

We might call the third strategy, ‘smokescreen’. Actors may create “shadow solutions” that present dangerous illusions of progress that serve mainly to distract attention and political will from more effective action. Elsewhere I argue that this is a sadly plausible concern about much of the history of international climate policy (Cf. Gardiner 2004; 2011a).

#### EVASION

The fourth area where the delegated responsibility model has explanatory value is when it comes to what we might call ‘the problem of evasion’, where existing actors aim to make certain claims on them disappear. For one thing, they may manipulate L2 responsibility assignments so that they are simply absolved. A live example of this in the climate case would be the tendency for actors perpetually to push deadlines for action into the further political future, beyond the political lifetimes of the actors currently responsible (Cf. Gardiner 2011a, chapters 3-4). For another thing, existing actors may manipulate the basic understanding of L1 responsibility so that even the claim of responsibility disappears. One specific strategy that seems highly effective is that of acting as if L2 assignments (including categories and agents) are fundamental, and thereby rendering L1 responsibility invisible. A more common strategy perhaps is for states in particular to oscillate between (a) invoking a conventional DRM for global affairs quite generally, (b) protesting that they are trying but defeated by the bad intentions of other L2 actors, yet (c) arguing that they do not have a specific L2 responsibility in this area, and are not particularly capable of solving the problem. In this “cunning wheeze”, most (or at least most powerful existing L2 agents) bemoan the general situation, blame everyone else, and quietly enjoy the benefits of the status quo. No one takes the *vital steps* of loudly (d) urging that the key issue might be the need for new institutions, or (e) insisting on the persistence of our L1 responsibilities in the absence of such institutions.

## DYSTOPIA

Fifth, the delegated responsibility model also has explanatory value in understanding wider dystopia. A natural objection to the idea of lingering L<sub>1</sub> responsibility at the individual level is that it seems deeply unrealistic and unfair in the world as we currently find it. Most people living ordinary lives are likely to protest: “How can it be *my* fault? In our world, I cannot be said to have any individual authority that is delegated or even respected. Others have the power, not me.” Jamieson himself may make a robust version of this claim when he says in passing: “[H]uman action is the driver, but it seems that things, not people, are in control ... Our corporations, governments, technologies, institutions and economic systems seem to have lives of their own” (Jamieson 2014, p. 1).

The DRM can explain the force of this alienation of the individual from responsibility at L<sub>1</sub>: other actors (L<sub>1</sub> and/or L<sub>2</sub>) have inhibited the individual’s ability to discharge his or her normal L<sub>1</sub> responsibilities, through blocking L<sub>2</sub> delegation for that agent, perhaps in multiple areas. In such cases, the fact that many individuals leading ordinary lives feel powerless may be no accident. In extreme scenarios, others have so arranged the social world that they have usurped power and thereby effective responsibility for these issues from such individuals. On my agency responsibility account (as on many others), this counts as a *central moral and political wrong*. Notably, the DRM does a reasonable job of describing how that wrong can be understood.

Our discussion so far has suggested that the DRM has considerable explanatory power, both in theoretical terms and when it comes to the real world example of climate change. It remains to ask what the implications of this discussion are for what is to be done.

## IV. A GLOBAL CONSTITUTIONAL CONVENTION

We have identified three major roots of the tyranny of the contemporary: ruthlessness, shallowness and obliviousness. The third and in my view most serious – obliviousness – rests on an *institutional gap* when it comes to registering and acting on intergenerational concern, and possibly also an *implicit hostility* of existing institutions designed with very different and conflicting aims. The DRM delivers an initial account of how we should understand our responsibilities in light of this gap, and the



agency responsibility view provides one kind of more specific understanding of the structure of those responsibilities.

A natural upshot of this diagnosis of the problem is that individuals retain L<sub>1</sub> (origin level) responsibilities to come together to support (i.e., create and sustain<sup>19</sup>) new L<sub>2</sub> (delegated level) institutions that will effectively fill the institutional gap, and do so in a way that is appropriately integrated with other aspects of agency (such as institutions with different L<sub>2</sub> responsibilities (or higher) as well as other individual responsibilities at L<sub>1</sub>). Questions then arise as to what the new institutions would look like, how they would fit into such a revised system of agency, and how to develop an achievable path towards both. Such questions would be demanding enough even if we had robust theories to guide us. However, our task is made more difficult by the fact that this is not so: we do not know in advance precisely what the best, or even adequate, versions of appropriate institutions or a revised system of agency would look like. Instead, we must proceed largely without such knowledge. Rather than straightforwardly applying robust theories, we must instead forge an ethics of the transition that moves us towards a better world without knowing in advance the precise destination (Gardiner 2011a).

#### UNDERSTANDING THE PROPOSAL

In this context, my proposal is that we should initiate a call for a global constitutional convention focused on future generations.<sup>20</sup> Let me explain. By ‘we’ here, I mean existing agents with responsibilities at L<sub>1</sub>, L<sub>2</sub> and other levels, including concerned individuals, interested community groups, national governments and transnational organizations. By ‘initiating a call’, I intend the procedural suggestion that we should (a) provoke and promote a call to action, that (b) seeks to engage a range of actors, that is based on (c) a claim that they have or should take on a set of responsibilities, and (d) a view about how to go about discharging those responsibilities. By “a global constitutional convention”, I have in mind a constitutional convention akin to that convened in Philadelphia in 1787 “to address the problems of the weak central government that existed under the Articles of Confederation”, and which led to the

19. Among other things, the reference to sustaining makes clear that there may, and in most cases are likely to be, further and residual responsibilities for individuals even given the delegation to new institutions.

20. The next three paragraphs draw on Gardiner 2014. Other proposals for intergenerational institutions tend to be focused on nation-states (e.g., Read 2012; Gonzalez-Ricoy and Gosseries 2016).



establishment of the US Constitution and the structure of federal government that persists to this day.<sup>21</sup> Our global constitutional convention would be convened to address the institutional gap surrounding concern for the future. Insofar as possible and prudent, it would do so without unnecessarily prejudging the outcome, whether in form or content.

The key components of my convention proposal can be understood in standard ways, as follows. First, I understand a *convention* to be “a representative body called together for some occasional or temporary purpose” and “constituted by statute to represent the people in their primary relations”. Second, the purpose of the convention is to establish a *constitutional system* in the minimal sense of “a set of norms (rules, principles or values) creating, structuring, and possibly defining the limits of, government power or authority.” Third, the *instigating* role of the convention would be to discuss, develop, make recommendations toward, and set in motion a process for the establishment of such a constitution. Fourth, the convention would take as its *primary subject matter* the need to adequately reflect and embody intergenerational concern, including through the protection of future generations, the promotion of their interests (e.g., their rights, claims, welfare, etc.), the discharging of earlier generations’ duties with respect to them, and so on.

I take this proposal to have several attractive features. To begin with, it is based in a deep political reality. It responds to the seriousness of climate change, and to persistent political inertia surrounding more modest initiatives. It calls attention to the heart of the problem (e.g., the failures of the current system, the need for an alternative, and the background issue of responsibility). It acknowledges that climate is only one instance of the tyranny of the contemporary, and we should expect others over the coming decades and centuries.<sup>22</sup> Moreover, though the proposal is ambitious, it is not alienating. Not only does it not succumb to despair, but it also does not needlessly polarize from the outset (e.g., by leaping to specific recommendations about filling the institutional gap). In addition, it acknowledges the existence of fundamental difficulties and anxieties, but uses them to start the right kind of debate, rather than to foreclose it. As such, in my view it is a promising candidate to serve as the subject of a wide, overlapping political consensus, at least among those with intergenerational

21. <https://history.state.gov/milestones/1784-1800/convention-and-ratification>

22. Within the perfect moral storm analysis, my focus is on global, intergenerational and ecological instantiations of the storm. However, the intergenerational aspect, represented here through the tyranny of the contemporary, can also manifest itself independently of these other features of the perfect moral storm.

concern. We might then turn to more specific conceptions of how to develop the proposal for a global constitutional convention that seek to support, refine and build on such a consensus. I have more to say about that project elsewhere.<sup>23</sup> For now, in order to clarify the basic proposal, let us turn to three initial objections.

#### WORLD GOVERNMENT?

One obvious initial objection to the proposal is that it is an overt appeal for an extremely powerful, highly centralized and overarching “world government” that would effectively subsume all other forms of government. Many appear to find the idea of this kind of centralized authority disturbing and perhaps outright offensive. However, whatever one thinks on this question, it is a misreading of my proposal to think that it mandates “world government” understood in this way.

The purpose of the convention is to consider how to establish “*a set of norms* (rules, principles or values) creating, structuring, and possibly defining the limits of, government power or authority.” The proposal for a global constitutional convention itself aims, as far as is possible and prudent, not to prejudge the outcome of that deliberation. It certainly does not presuppose that a highly-centralized world government would emerge. Indeed, it envisions that a central question facing the convention itself will be how to develop a broader system of institutions and practices than is currently in place, that will, on the one hand, be capable of and likely to produce effective action (and so reflect some desirable features of a powerful and highly centralized global authority), but also, on the other hand, allow for the significant preservation of existing institutions to serve as a bulwark against the excesses of any newly created ones (e.g., by neutralizing the standing threats posed by any new bodies, through methods such as the traditional separation of powers, “anti-power”, etc.). How this question is answered is not seriously prejudged merely by proposing a global constitutional convention to address it.

#### UNDEMOCRATIC?

A second initial objection is that a global constitutional convention and any institutions that might flow from it would be inherently undemocratic (or, more neutrally, politically illegitimate or unrepresentative), especially because they would un-

23. See Gardiner (2014) for discussion of guidelines for the global constitutional convention.

dermine the sovereignty of conventional nation states, and thereby obstruct the will of current peoples.

As a theoretical objection, this worry strikes me as misguided. In particular, it appears to get things backwards. The motivating idea for the global constitutional convention is that *there is an existing institutional gap that itself leads to an inappropriate and undemocratic global system*. Consider three essential points.

First, according to my diagnosis, people's legitimate intergenerational concerns are currently not being appropriately signaled or represented by current institutions. Therefore, the DRM implies that the current system is defective and that we have a responsibility to address this defect.

Second, part of the problem is that current institutions, designed for other (L<sub>2</sub>) purposes, risk *usurping* the role that should be played by good intergenerational institutions. They therefore obstruct effective responsibility assignment (from L<sub>1</sub> to L<sub>2</sub>) by individuals, impede the moral project of responsible agency for such individuals, and thereby undermine their own (L<sub>2</sub>) authority. Consequently, *the real threat to democratic values here is that current institutions threaten to confound the "will of the people" properly understood*. As a result, and contrary to the initial objection, successfully filling the institutional gap (with new or adjusted L<sub>2</sub> institutions) would *improve* the political legitimacy of the overall institutional order. It would pay due respect to concerns that current (as well as past and future) people have but which are underappreciated and often thwarted by existing arrangements. In particular, to the extent that the sovereignty of nation states as currently organized illegitimately impedes such concerns, some adjustment to that sovereignty may be necessary.

Third, more generally, the objection requires further elaboration. One concern is that it presupposes a monolithic view of sovereignty – that the people must have just one, all-powerful representative (or L<sub>2</sub> agent), and this must be the nation state—that is unattractive and runs counter to many existing systems and practices. Though neither the DRM nor the global constitutional convention proposal rule out such a monolithic structure, they also do not assume that one is required. Instead, they allow for the possibility of more pluralistic models, whereby individuals may delegate their L<sub>1</sub> responsibilities to a number of institutions at L<sub>2</sub> (or higher). These may be distinct entities, or overlapping, or related in other complicated ways. They may be further unified in a monolithic overarching structure, but they may not. Importantly, often there are reasons to discourage too much unity, such as when there is a need for checks and balances. For example, federal systems such as the United States divide

responsibilities within the federal government, and also between the federal government and state and local authorities. *Each institution in the system is seen as a legitimate representative of the people*, often responsible for distinct but overlapping L2 responsibilities. None is seen as compromising the “sovereignty of the people” simply in virtue of its existence, nor merely because it may sometimes come into conflict with other L2 institutions. Instead, each institution represents some legitimate aspect of the will of the people (and their L1 responsibilities), and has its own legitimate sphere of influence, including when that sphere overlaps with the spheres of other L2 institutions.

### UNFEASIBLE?

A third initial objection is that the proposal for a global constitutional convention is (perhaps laughably) unfeasible given the world as we currently know it, and especially on a time frame relevant to successfully addressing a problem like climate change. I have a number of responses.

Let us begin with the specific concern about the time-frame. First, it is worth noting that, as serious as the climate problem is, the proposal for a global constitutional convention is aimed not just, or even primarily, at climate change, but at a much wider challenge of which climate change is but one instance (i.e., the tyranny of the contemporary in particular, and the perfect moral storm more generally). Hence, the proposal would have an important purpose even if it did not contribute to solving the climate problem, but instead helped to forestall further intergenerational problems in the future.

Second, the proposal is not intrinsically antagonistic to other processes that might help in the climate (or other) cases, and may even facilitate such processes. For instance, enthusiasts for the current Paris Agreement may still continue to support the existing UNFCCC process as part of a two-track approach, where one track is the UNFCCC and the other is the global constitutional convention. Indeed, in my view progress on the global constitutional convention promises to make substantial action within the UNFCCC process more likely and more robust.

Third, we should be evenhanded and compare proposals on a level playing field. In particular, if the timeframe is an issue, one should not forget that the main rivals to the proposal for a global constitutional convention have not so far proved quick at producing action. In the climate case, the UNFCCC process has already taken more than 20 years. (Recall Mary Robinson’s observation that “global negotiations

have proceeded at a glacial pace”). Moreover, many continue to criticize the current ambition and structure of the Paris agreement, calling its approach much too modest, and pressing for further evolution over the next decade or so (Cf. Oxfam et al. 2015; Anderson and Peters 2016). At the time of writing (November 2016), the recent election of the Trump Administration in the United States has underlined the fragility of the Paris approach.

Still, fourth, and most importantly, we should question the implicit assumption that the call for a global constitutional convention necessarily involves a commitment to a very long process. For instance, the US constitutional convention took place over only a few months, and its recommendations were ratified and implemented within a few years. With real political will, and the sense that a global constitutional convention may actually produce effective institutions, similarly fast results may be achieved even allowing for the necessary due deliberation.

This brings us to broader issues of feasibility. Again, I will make just a few points. First, we should recognize that notions of political feasibility are often highly treacherous. In my lifetime, many things said to be politically infeasible (the fall of the Berlin Wall, the peaceful collapse of Apartheid in South Africa, the election of the first black US President) have come to pass, and on a timescale far shorter than mainstream understandings of political reality would allow. As some people like to say, sometimes one reaches social and political “tipping points”. After decades of relative inaction, we may be getting closer to one on climate and this may facilitate the proposal for a global constitutional convention. This may be so even if in the short-term the prospects for robust action become more bleak. Such deadlock may make the institutional problem even clearer.

Second, evenhandedness in comparing proposals is an issue here as well. For instance, the main alternative proposal to emerge recently is that of geoengineering—grand technological interventions into the global climate system, such as stratospheric sulfate injection (‘SSI’). Notably, the pursuit of SSI is often motivated under the idea that a back up plan or ‘Plan B’ is needed in case conventional climate negotiations fail, or do not deliver action quickly enough (Fragniere and Gardiner 2016). Yet the ‘back up plan’ rationale seems to provide at least as much reason to push for a global constitutional convention, especially as it is arguably more socially and technically achievable.

To push this point home, we might note that the ‘back up plan’ rhetoric for geo-engineering remains curiously popular even when it comes to promoting “magical”

technologies: technologies that do not yet exist, and are not very close to existing, especially on the scale needed. Some critics, for example, complain that current IPCC pathways towards 2 degrees Celsius “require massive deployment of Biomass Energy with Carbon Capture and Storage (BECCS)”, when there is “a distinct lack of evidence to determine whether BECCS is technically feasible, economically affordable, environmentally benign, socially acceptable and politically viable at a material scale” (Kruger et al. 2016). Moreover, elsewhere I argue that there is no escape from the institutional questions: even minimally decent forms of geoengineering (and also many indecent but perhaps morally tolerable forms) will also require significant institutional reform of the kind that the global constitutional convention is intended to address (Gardiner 2011a).

The third point about broad feasibility is that the proposal for a global constitutional convention at least has the virtues of highlighting the depths of the current structural problems, challenging us to confront them directly in an appropriately serious venue, and yet being comparatively open and initially minimalist about the precise content of what should emerge. Moreover, critically, the real obstacles that would need to be overcome are those likely to arise for *any* proposal that may *actually make a crucial difference*. Consequently, while we should acknowledge that deep pessimism about the proposal for a global constitutional convention is easy to understand, especially in the face of the tyranny of the contemporary and other lurking defects of current institutions, we cannot afford to give in to it. The key issues highlighted by the proposal for a global constitutional convention are *just those that must be addressed* if we are to confront the challenge that faces us, and so discharge our responsibilities.

## V. CONCLUSION

I have argued that we face an ongoing failure of the global political system in dealing with paradigmatically global, intergenerational and ecological problems, such as climate change. Existing institutions tend to crowd out intergenerational concern, and thereby encourage the live threats of a tyranny of the contemporary and a corruption of the discourse. Though such threats have several sources, I focused on the institutional gap around the issue of registering intergenerational concern.

My general approach has been to say that meeting these threats will require both accepting responsibility for the future and filling in the institutional gap. In order to understand the first (accepting responsibility), I proposed a traditional delegated



responsibility model of the transmission of individual responsibility to collectives. I then developed the delegated responsibility model by suggesting how it leads us to understand both past failures and prospective responsibility. For the second (filling the institutional gap), I proposed that the current generation should take responsibility by initiating a call for a global constitutional convention focused on concern for future generations. I then briefly defended this proposal, in part by rejecting three initial objections, that the proposal is an overt call for heavily centralized world government, that it is inherently undemocratic, and that it is unfeasible. Though my proposal is only a first step and requires further discussion, my current aim is simply to open that important debate.<sup>24</sup>

*Acknowledgements: This paper brings together, and substantially extends, arguments I have given elsewhere. Previous versions were presented at conferences at the Carnegie Council for Ethics in International Affairs (New York), the Institute for Futures Studies (Stockholm), the University of Leeds, and University College, Dublin. I am grateful to the audiences for their comments, and especially to Gustav Arrhenius, Simon Caney, Alix Dietzel, Elizabeth Finneron-Burns, Guy Kahane, Rob Lawlor, Alex Lenferna, Christopher Megone, Philip Pettit, Julian Savulescu, Paul Tubig, Alexa Zellentin and two anonymous reviewers. I also thank Ken Shockley for early discussion at COP 21 in Durban (2010).*

#### REFERENCES

Anderson, Kevin and Glen Peters. 2016. 'The Trouble with Negative Emissions', *Science* 354 (6309), 182-183.

Bertram, Christopher. 2009. 'Exploitation and Future Generations'. In Axel Gosseries and Lukas Meyer, eds. *Intergenerational Justice*. Oxford.

Fragniere, Augustin and Stephen M. Gardiner. 2016. 'Why Geoengineering is not Plan B'. In Christopher Preston, ed. *Justice and Geoengineering*. Rowman and Littlefield.

Gardiner, Stephen M. 2017a. 'The Threat of Intergenerational Extortion: On the Temptation to Become the Climate Mafia, Masquerading as an Intergenerational Robin Hood', *Canadian Journal of Philosophy*. Special issue on intergenerational ethics, edited by Rahul Kumar.

———. 2017b. 'Climate Ethics in a Dark and Dangerous Time', *Ethics* 127: 430-465.

24. Elsewhere I take a second step by advancing guidelines concerning the characteristics, aims, composition and scope of the convention (Gardiner 2014).



———. 2015. 'The Need for a Public "Explosion" in the Principles of Nuclear Ethics'. In Sabine Roeser and Behnam Taebi, eds. *The Ethics of Nuclear Energy: Risk, Justice and Democracy in the post-Fukushima Era*. Cambridge.

———. 2014. 'A Call for a Global Constitutional Convention focused on Future Generations', *Ethics and International Affairs* 28.3, 299-315.

———. 2013. 'Reflecting on A Perfect Moral Storm: Replies to Di Paola, Jamieson, Pellegrino, Preston, Sandler and Smead', *Philosophy and Public Issues (New Series)* 3.1, 89-134.

———. 2011a. *A Perfect Moral Storm: The Ethical Challenge of Climate Change*. Oxford.

———. 2011b. 'Rawls and Climate Change: Does Rawlsian Political Philosophy Pass the Global Test?', *Critical Review of International Social and Political Philosophy* 14.2, 125-151. Special issue, also appearing as: Gideon Calder and Catriona McKinnon, eds. 2011. *Climate Change and Liberal Priorities*. Routledge.

———. 2011c. 'Is No One Responsible for Global Environmental Tragedy? Climate Change as a Challenge to Our Ethical Concepts'. In Denis Arnold, ed., *Ethics and Global Climate Change*. Cambridge, 38-59.

———. 2004. 'The Global Warming Tragedy and the Dangerous Illusion of the Kyoto Protocol', *Ethics and International Affairs* 18.1, 23-39.

———. Unpublished. 'If a Climate Catastrophe is Possible, is Everything Permitted?' Paper presented at University of Melbourne, 2014.

Gardiner, Stephen M. and David Weisbach. 2016. *Debating Climate Ethics*. Oxford.

Gonzalez-Ricoy, Inigo and Axel Gosseries. Eds. 2016. *Institutions for Future Generations*. Oxford.

Jamieson, Dale. 2014. *Reason in a Dark Time*. Oxford.

———. 2013. 'Jack, Jill and Jane in a Perfect Moral Storm'. *Philosophy and Public Issues (New Series)* 3.1.

———. 2010. 'Climate Change, Responsibility and Justice', *Science and Engineering Ethics* 16, 431-445

———. 1992. 'Ethics, Public Policy and Global Warming', *Science, Technology, and Human Values* 17, 139-153.

Kruger, Tim, Oliver Geden and Steve Rayner. 2016. 'Abandon Hype in Climate Models'. *Guardian*, April 26.

Milman, Oliver. 2015. "James Hansen, Father of Climate Change Awareness, Calls Paris Talks a 'Fraud'", *Guardian*, December 12.

Morris, Gouverneur. 1799. 'Oration Upon Washington'. In Max Farrand, ed. *The Records of the Federal Convention of 1787*, Volume 3, 382.

Nolt, John. 2011. 'Greenhouse Gas Emission and the Domination of Posterity'. In Denis Arnold, ed., *The Ethics of Global Climate Change*, Cambridge, 60-76.

Oxfam et al. 2015. *Fair Shares: A Civil Society Equity Review of INDCs*, report (Oxford: Oxfam International). Available at: [https://www.oxfam.org/sites/www.oxfam.org/files/file\\_attachments/ib-civil-society-review-climate-indcs-191015-en\\_2.pdf](https://www.oxfam.org/sites/www.oxfam.org/files/file_attachments/ib-civil-society-review-climate-indcs-191015-en_2.pdf).

Read, Rupert. 2012. *Guardians of the Future: A Constitutional Case for Representing and Protecting Future People* (Weymouth, UK: Green House).

Robinson, Mary. 2008. 'Foreword'. In *Climate Change and Human Rights: A Rough Guide*. International Council on Human Rights, pp. iii-iv.

Sinnott-Armstrong, Walter. 'It's Not My Fault'. In Walter Sinnott-Armstrong and Richard Howarth, eds. *Perspectives on Climate Change*. New York: Elsevier, 221-253.

Smith, Patrick Taylor. 'The Intergenerational Storm: Dilemma or Domination', *Philosophy and Public Issues (New Series)* 3.1.

Von Bulow, Michael. 2009. 'Failure in Copenhagen is Not an Option', COP-Web Site <http://en.cop15.dk/news/view+news?newsid=2257> (October 2). [no longer available]

Young, Iris Marion. 2013. *Responsibility for Justice*. Oxford.

# The Death Penalty Debate: Four Problems and New Philosophical Perspectives

Masaki Ichinose

*The University of Tokyo*

## ABSTRACT

This paper aims at bringing a new philosophical perspective to the current debate on the death penalty through a discussion of peculiar kinds of uncertainties that surround the death penalty. I focus on laying out the philosophical argument, with the aim of stimulating and restructuring the death penalty debate.

I will begin by describing views about punishment that argue in favour of either retaining the death penalty ('retentionism') or abolishing it ('abolitionism'). I will then argue that we should not ignore the so-called "whom-question", i.e. "To whom should we justify the system of punishment?" I identify three distinct chronological stages to address this problem, namely, "the Harm Stage", "the Blame Stage", and "the Danger Stage".

I will also identify four problems arising from specific kinds of uncertainties present in current death penalty debates: (1) uncertainty in harm, (2) uncertainty in blame, (3) uncertainty in rights, and (4) uncertainty in causal consequences. In the course of examining these four problems, I will propose an 'impossibilist' position towards the death penalty, according to which the notion of the death penalty is inherently contradictory.

Finally, I will suggest that it may be possible to apply this philosophical perspective to the justice system more broadly, in particular to the maximalist approach to restorative justice.

---

## 1. TO WHOM SHOULD PUNISHMENT BE JUSTIFIED?

What, exactly, are we doing when we justify a system of punishment? The process of justifying something is intrinsically connected with the process of persuading someone to accept it. When we justify a certain belief, our aim is to demonstrate reasonable grounds for people to believe it. Likewise, when we justify a system of taxation, we intend to demonstrate the necessity and fairness of the system to taxpayers.

What, then, are we justifying when we justify a system of punishment? To whom should we provide legitimate reasons for the system? It is easy to understand to whom we justify punishment when that punishment is administered by, for example, charging a fine. In this case, we persuade violators to pay the fine by bringing to their attention the harm that they have caused, harm which needs to be compensated. (Please note that I am only mentioning the primitive basis of the process of justification.) While we often generalise this process to include people in general or society as a whole, the process of justification would not work without convincing the people who are directly concerned (in this case, violators), at least theoretically, that this is a justified punishment, despite their subjective objections or psychological opposition. We could paraphrase this point per Scanlon's 'idea of a justification which it would be unreasonable to reject' (1982, p.117). That is to say, in justifying the application of the system of punishment, we should satisfy the condition that each person concerned (especially the violator) is aware of having no grounds to reasonably reject the application of the system, even if they do in fact reject it from their personal, self-interested point of view.

In fact, if the violator is not theoretically persuaded at all in any sense—that is, if they cannot understand the justification as a justification—we must consider the possibility that they suffer some disorder or disability that affects their criminal responsibility.

We should also take into account the case of some extreme and fanatical terrorists. They might not understand the physical treatment inflicted on them in the name of punishment as a punishment at all. Rather, they might interpret their being physically harmed as an admirable result of their heroic behaviour. The notion of

punishment is not easily applied to these cases, where the use of physical restraint is more like that applied to wild animals. Punishment can be successful only if those who are punished understand the event as punishment.

This line of argument entirely conforms to the traditional context in philosophy concerning the concept of a “person”, who is regarded as the moral and legal agent responsible for his or her actions, including crimes. John Locke, a 17th-century English philosopher, introduced and established this concept, basing it on ‘consciousness’. According to Locke, a person ‘is a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness’ (1975, Book 2, Chapter 27, Section 9). This suggests that moral or legal punishments for the person should be accompanied by consciousnesses (in a Lockean sense) of the agent. In other words, when punishment is legally imposed on someone, the person to be punished must be conscious of the punishment as a punishment; that is, the person should understand the event as a justified imposition of some harm.<sup>1</sup>

However, there is a problem here, which arises in particular for the death penalty but not for other kinds of punishment. The question that I raise here is ‘to whom do we justify the death penalty?’ People might say it should be justified to society, as the death penalty is one of the social institutions to which we consent, whether explicitly or tacitly. This is true. However, if my claims above about justification are correct, the justification of the death penalty must involve the condemned convict coming to understand the justification at least at a theoretical level. Otherwise, to be executed would not be considered a punishment but rather something akin to the extermination of a dangerous animal. The question I want to focus on in particular is this: should this justification be provided *before* administering capital punishment or *whilst* administering capital punishment?

1. Strangely, few Locke scholars have seriously tried to understand the Lockean meaning of punishment, which is developed in his *Second Treatise* (Locke 1960), in the light of his theory of personal identity based upon ‘consciousness’, which is discussed in his *Essay Concerning Human Understanding*. Taking into account the fact that ‘person’ appears as the key word in both works of Locke, we must bridge the gap between his two works by rethinking the universal significance of ‘person’ in his arguments. There were, however, some controversies concerning how Locke evaluates the death penalty. See Calvert (1993) and Simmons (1994).

## 2. 'IMPOSSIBILISM'

Generally, in order for the justification of punishment to work, it is necessary for convicts to understand that this is a punishment before it is carried out and that they cannot reasonably reject the justification, regardless of any personal objection they may have. However, that is not sufficient, because if they do not understand *at the moment of execution* that something harmful being inflicted is a punishment, then its being inflicted would simply result in mere physical harm rather than an institutional response based on theoretical justification. The justification for punishment must be, at least theoretically, accepted *both before and during* its application.<sup>2</sup> This requirement can be achieved with regard to many types of punishment, such as fines or imprisonment. However, the situation is radically different in the case of the death penalty, for in this case, when it is carried out, the convict, by definition, disappears. During and (in the absence of an afterlife) after the punishment, the convict cannot understand the nature and justification of the punishment. Can we say then that this is a punishment? This is a question which deserves further thought.

On the one hand, the death penalty, once executed, logically implies the nonexistence of the person punished; therefore, by definition, that person will not be conscious of being punished at the moment of execution. However, punishment must be accompanied by the convict's consciousness or understanding of the significance of the punishment, as far as we accept the traditional concept of the person as a moral and legal agent upon whom punishment could be imposed. It may be suggested that everything leading up to the execution—being on death row, entering the execution chamber, being strapped down—is a kind of punishment that the convict is conscious of and is qualitatively different from mere incarceration. However, those phases are factors merely concomitant with the death penalty. The core essence of being executed lies in being killed or dying. Therefore, if the phases of anticipation were to occur but finally the convict were not killed, the death penalty would not have been carried out. The death penalty logically results in the convict's *not being conscious* of being executed, and yet, for it to be a punishment, the death penalty requires the convict to

2. There is an additional question about whether justification is needed *after* the execution when the convict is no longer around, in addition to 'before' and 'during'. According to my understanding of justification, the process of justification must begin with making each person concerned understand what there is no reason to reject, but that is just a starting, necessary point. Justification must go beyond the initial phase to acquiring general consent from society. In this sense, justification seems to be needed even 'after' the execution. Actually, if there is no need for justification after the execution, that sounds less like punishment based on a system of justice than merely physical disposal.

be conscious of being executed. We could notate this in the form of conjunction in the following way in order to make my point as clear as possible:

~ PCE & PCE

(PCE: 'the person is conscious of being executed under the name of punishment')

If this is correct, then we must conclude that the concept of the death penalty is a manifest contradiction in terms. In other words, the death penalty should be regarded as conceptually impossible, even before we take part in longstanding debates between retentionism and abolitionism. This purely philosophical view of the death penalty could be called 'impossibilism' (i.e. the death penalty is conceptually impossible), and could be classified as a third possible view on the death penalty, distinct from retentionism and abolitionism. A naïve objection against this impossibilist view might counter that the death penalty is actually carried out in some countries so that it is not impossible but obviously possible. The impossibilist answer to this objection is that, based on a coherent sense of what it means for a punishment to be justified, that execution in such countries is not the death *penalty* but rather unjustified lethal physical *violence*.

I am not entirely certain whether the 'impossibilist' view would truly make sense in the light of the contemporary debates on the death penalty. These debates take place between two camps as I referred to above:

*Retentionism* (the death penalty should be retained): generally argued with reference to victims' feelings and the deterrence effects expected by execution.

*Abolitionism* (the death penalty should be abolished): generally argued through appeals to the cruelty of execution, the possibility of misjudgements in the trial etc.

The grounds mentioned by both camps are, theoretically speaking, applicable to punishment in general in addition to the death penalty specifically. I will mention those two camps later again in a more detailed way in order to make a contrast between standard debates and my own view. However, my argument above for 'impossibilism', does suggest that there is an uncertainty specific to the death penalty as opposed to other types of punishment. I believe that this uncertainty must be considered when we discuss the death penalty, at least from a philosophical perspective. Otherwise we may lose sight of what we are attempting to achieve.

A related idea to the 'impossibilism' of the death penalty may emerge, if we accept the fact that the death penalty is mainly imposed on those convicted of homi-



cide. This idea is related to the understanding of death proposed by Epicurus, who provides the following argument (Diogenes Laertius 1925, p. 650-1):

*Death, therefore, the most awful of evils, is nothing to us, seeing that, when we are, death is not come, and, when death is come, we are not. It is nothing, then, either to the living or to the dead, for with the living it is not and the dead exist no longer.*

We can call this Epicurean view ‘the harmlessness theory of death’ (HTD). If we accept HTD, it follows, quite surprisingly, that there is no direct victim in the case of homicide insofar as we define ‘victim’ to be a person who suffers harm as a result of a crime. For according to HTD, people who have been killed and are now dead suffer nothing—neither benefits nor harms—because, as they do not exist, they cannot be victims. If this is true, there is no victim in the case of homicide, and it must be unreasonable to impose what is supposed to be the ultimate punishment<sup>3</sup>—that is, the death penalty—on those offenders who have killed others.

This argument might sound utterly absurd, particularly if it is extended beyond offenders and victims to people in general, as one merit of the death penalty seems to lie in reducing people’s fear of death by homicide. However, although this argument from HTD might sound bizarre and counterintuitive, we should accept it at the theoretical level, to the extent that we find HTD valid.<sup>4</sup> Clearly, this argument, which is based on the nonexistence of victims, could logically lead to another possibilist argument concerning the death penalty.

There are many points to be more carefully examined regarding both types of ‘impossibilism’, which I will skip here. However, I must stop to ponder a natural reaction. My question above, ‘To whom do we justify?’, which introduced ‘impossibilism’,

3. Is it true that the death penalty is the ultimate punishment? Can we not suppose that the death penalty is less harmful than a life sentence or very lengthy incarceration? However, this view regarding the death penalty as less harmful than a lifelong sentence could lead to a paradox. If this order of severity as punishment is valid, it may be possible to reduce the lifelong sentence (due to an amnesty, some consideration on the prisoner’s rehabilitation, or something like that) to the death penalty. If this is the case, prisoners given the lifelong sentence will not make an effort at all to rehabilitate themselves, due to fear of the sentence being reduced to the death penalty. In addition, if a person is likely to be sentenced to death, the person might try to commit a more heinous crime, perhaps even in the court in order to be given a more severe sentence, i.e. a life sentence in prison. That is a paradox drawn from human nature.

4. On the current debates on ‘HTD’ of Epicurus, see Fischer (1993). Of course, there are lots of objections against the Epicurean view. The most typical objection is that death deprives people of their chance to enjoy life, and therefore death is harmful. However, it seems to me that “whom-question” must be raised again here. To whom is the deprivation of this chance harmful? In any case, the metaphysics of death is a popular topic in contemporary philosophy, which should involve not only metaphysical issues but also ethical and epistemological problems.

might sound eccentric, because, roughly speaking, theoretical arguments of justification are usually deployed in a generalised way and do not need to acknowledge who those arguments are directed at. Yet, I believe that this normal attitude towards justification is not always correct. Instead, our behaviour, when justifying something, focuses primarily on theoretically persuading those who are unwilling to accept the item being justified. If nobody refuses to accept it, then it is completely unnecessary to provide its justification. For instance, to use a common sense example, nobody doubts the existence of the earth. Therefore, nobody takes it to be necessary to justify the existence of the earth. Alternatively, a justification for keeping coal-fired power generation, the continued use of which is not universally accepted due to global warming, is deemed necessary. In other words, justification is not a procedure lacking a particular addressee, but an activity that addresses the particular person in a definite way, at least at first. In fact, it seems to me that the reason that current debates on the death penalty become deadlocked is that crucial distinctions are not appropriately made. I think that such a situation originates from not clearly asking to whom we are addressing our arguments, or whom we are discussing. As far as I know, there have been very few arguments within the death penalty debate that take into account the homicide victim, despite the victim's unique status in the issue. This is one example where the debate can be accused of ignoring the 'whom-question', so I will clarify this issue by adopting a strategy in which this 'whom-question' is addressed.

### 3. THREE CHRONOLOGICAL STAGES

Following my strategy, I will first introduce a distinction between three chronological stages in the death penalty. In order to make my argument as simple as possible, I will assume that the death penalty is imposed on those who have been convicted of homicide, although I acknowledge there are other crimes which could result in the death penalty. In that sense, the three stages of the death penalty correspond to the three distinct phases arising from homicide.

The first stage takes place at the time of killing; the fact that someone was killed must be highlighted. However, precisely what happened? If we accept the HTD, we should suppose that nothing harmful happened in the case of homicide. Although counterintuitive, let's see where this argument leads. However, first, I will acknowledge that we cannot cover all contexts concerning the justification of the death penalty by discussing whether or not killing harms the killed victim. Even if we

accept for argument's sake that homicide does not harm the victim, that is only part of the issue. Other people, particularly the bereaved families of those killed, are seriously harmed by homicide. More generally, society as a whole is harmed, as the fear of homicide becomes more widespread in society.

Moreover, our basic premise, HTD, is controversial. Whether HTD is convincing remains an unanswered question. There is still a very real possibility that those who were killed do suffer harm in a straightforward sense, which conforms to most people's strong intuition. In any event, we can call this first stage, the 'Harm Stage', because harm is what is most salient in this phase, either harm to the victims or others in society at large. If a justification for the death penalty is to take this Harm Stage seriously, the overwhelming focus must be on the direct victims themselves, who actually suffer the harm. This is the central core of the issue, as well as the starting point of all further problems.

The second stage appears after the killing. After a homicide, it is common to blame and to feel anger towards the perpetrator or perpetrators, and this can be described as a natural, moral, or emotional reaction. However, it is not proven that blaming or feeling angry is indeed natural, as it has not been proven that such feelings would arise irrespective of our cultural understanding of the social significance of killing. The phenomenon of blaming and the prevalence of anger when a homicide is committed could be a culture-laden phenomenon rather than a natural emotion. Nevertheless, many people actually do blame perpetrators or feel anger towards them for killing someone, and this is one of the basic ideas used to justify a system of 'retributive justice'. The core of retributive justice is that punishment should be imposed on the offenders themselves (rather than other people, such as the offenders' family). This retributive impulse seems to be the most fundamental basis of the system of punishment, even though we often also rely on some consequentialist justification for punishment (e.g. preventing someone from repeating an offence). In addition, offenders are the recipients of blame or anger from society, which suggests that blaming or expressing anger has a crucial function in retributive justice. I will call this second phase the 'Blame Stage', which extends to the period of the execution. Actually, the act of blaming seems to delineate what needs to be resolved in this phase. Attempting to justify the death penalty by acknowledging this Blame Stage (or retributive justification) in terms of proportionality is the most common strategy. That is to say, *lex talionis* applies here—'an eye for an eye'. This is the justification that not only considers people in general, including victims who blame perpetrators, but

also attempts to persuade perpetrators that this is retribution resulting from their own harmful behaviours.

The final stage in the process concerning the death penalty appears after the execution; in this stage, what matters most is how beneficial the execution is to society. Any system in our society must be considered in the light of its cost-effectiveness. This extends even to cultural or artistic institutions, although at first glance they seem to be far from producing any practical effects. In this context, benefits are interpreted quite broadly; creating intellectual satisfaction, for example, is counted as a benefit. Clearly, this is a utilitarian standpoint. We can apply this view to the system of punishment, or the death penalty, if it is accepted. That is, the death penalty may be justified if its benefits to society are higher than its costs. What, then, are the costs, and what are the benefits? Obviously, we must consider basic expenses, such as the maintenance and labour costs of the institution keeping the prisoner on death row. However, in the case of the death penalty, there is a special cost to be considered, namely, the emotional reaction of people in society in response to killing humans, even when officially sanctioned as a punishment. Some feel that it is cruel to kill a person, regardless of the reason.

On the other hand, what is the expected benefit of the death penalty? The 'deterrent effect' is usually mentioned as a benefit that the death penalty can bring about in the future. In that case, what needs to be shown if we are to draw analogies with the previous two stages? When people try to justify the death penalty by mentioning its deterrent effect, they seem to be comparing a society without the death penalty to one with the death penalty. Then they argue that citizens in a society with the death penalty are at less risk of being killed or seriously victimised than those in a society without the death penalty. In other words, the death penalty could reduce the danger of being killed or seriously victimised in the future. Therefore, we could call this third phase the 'Danger Stage'. In this stage, we focus on the danger that might affect people in the future, including future generations. This is a radically different circumstance from those of the previous two stages in that the Danger Stage targets people who have nothing to do with a particular homicide.

#### 4. ANALOGY FROM NATURAL DISASTERS

The three chronological stages that I have presented in relation to the death penalty are found in other types of punishment as well. Initially, any punishment must

stem from some level of harm (including harm to the law), and this is a *sine qua non* for the issue of punishment to arise. Blaming and its retributive reaction must follow that harm, and subsequently some social deterrent is expected to result. However, we should carefully distinguish between the death penalty and other forms of punishment. With other forms of punishment, direct victims undoubtedly exist, and those convicted of harming such victims are aware they are being punished. In addition, rehabilitating perpetrators in order for them to return to society—one aspect of the deterrent effect—can work in principle. However, this aspect of deterrence cannot apply to the death penalty because executed criminals cannot be aware of being punished by definition, and the notion of rehabilitation does not make sense by definition. Only this quite obvious observation can clarify that there is a crucial, intrinsic difference or distinction between the death penalty and other forms of punishment. Theories about the death penalty must seriously consider this difference; we cannot rely on theories that treat the death penalty on a par with other forms of punishment.

Moreover, the three chronological stages that have been introduced above are fundamentally different from each other. In reality, the subjects or people that we discuss and on whom we focus are different from stage to stage. In this respect, one of my points in this article is to underline the crucial need to discuss the issues of the death penalty by drawing a clear distinction between those stages. I am not claiming that only one of those stages is important. I am aware that each stage has its own significance; therefore, we should consider all three. However, we should be conscious of the distinctions when discussing the death penalty.

To make my point more understandable, I will suggest an analogy with natural disasters. Specifically, I will use as an analogy the biggest earthquake in Japan in the past millennium—the quake of 11 March 2011 (hereafter the 2011 quake). Of course, at first glance, earthquakes are substantially different from homicides. However, there is a close similarity between the 2011 quake and homicides, because although most of the harm that occurred was due to the earthquake and tsunami, in fact people were also harmed and killed during the 2011 quake at least partially due to human errors, such as the failure of the government's policy on tsunamis and nuclear power plants. Thus, it is quite easy in the case of the 2011 quake to distinguish between three aspects, all of which are different from each other.

(1) We must recognise victims who were killed in the tsunami or suffered hard-

ship at shelters.<sup>5</sup> This is the core as well as the starting point of all problems. What matters here is rescuing victims, and expressing our condolences.

(2) Then we will consider victims and people in general who hold the government and the nuclear power company responsible for political and technical mistakes. What usually matters here is the issue of responsibility and compensation.

(3) Finally, we can consider people's interests in improving preventive measures taken to reduce damages by tsunami and nuclear-plant-related accidents in the future. What matters in this context is the reduction of danger in the future by learning from the 2011 quake.

Nobody will fail to notice that these three aspects are three completely different issues, which can be seen in exactly the same manner in the case of the death penalty. Aspects (1), (2), and (3) correspond respectively to the Harm Stage, the Blame Stage, and the Danger Stage. Undoubtedly, none of these three aspects should be ignored and they actually appear in a mutually intertwined manner: the more successful the preventive measures are, the fewer victims will be produced by tsunami and nuclear-plant accidents in the future. Those aspects affect each other. Likewise, we must consider each of the three stages regarding the death penalty.

## 5. INITIAL HARM

The arguments thus far provide the basic standpoint that I want to propose concerning the debates on the death penalty. I want to investigate the issue of the death penalty by sharply distinguishing between these three stages and by simultaneously considering them all equally. By following this strategy, I will demonstrate that there are intrinsic uncertainties, and four problems resulting from those uncertainties, in the system of the death penalty. In so doing I will raise a novel objection to the contemporary debate over the death penalty.

Roughly speaking, as I have previously mentioned, the death penalty debate continues to involve the two opposing views of abolitionism and retentionism (or perhaps, in the case of abolitionist countries, revivalism). It seems that the main argu-

5. In fact, the hardships suffered by those forced to flee to shelters constituted the main problem resulting from the nuclear power plants accident. In general, radiation exposure is the most well-known problem arising from nuclear power plant accidents, but it is not always the case. In particular in the case of the Fukushima nuclear power plant accident in Japan, the overestimation of the danger of radiation exposure, and evacuation activities resulting from that overestimation, caused the biggest and the most serious problems including many of the deaths. We always have to take the risk-tradeoff into account. Radiation exposure is just one risk, and is not the only risk to be considered. See Ichinose (2016).



ments to support or justify each of the two traditional views (which I have briefly described in section 2 above) have already been exhausted. What matters in this context is whether the death penalty can be justified, and then whether—if it is justifiable—it should be justified in terms of retributivism or utilitarianism. That is the standard way of the debate on the death penalty. For example, when the retributive standpoint is used to justify the death penalty, the notion of proportionality as an element of fairness or social justice might be relevant, apart from the issue of whether proportionality should be measured cardinally or ordinally (see von Hirsch 1993, pp. 6–19). In other words, if one person has killed another, then that person too ought to be killed—that is, executed—in order to achieve fairness. However, as other scholars such as Tonry (1994) have argued, it is rather problematic to apply the notion of proportionality to the practice of punishment because it seems that there is no objective measure of offence, culpability, or responsibility. Rather, the notion of parsimony<sup>6</sup> is often mentioned in these contexts as a more practical and fairer principle than the notion of proportionality.

However, according to my argument above, such debates are inadequate if they are simply applied to the case of the death penalty. Proportionality between which two things is being discussed? Most likely, what is considered here is the proportionality between harm by homicide (where the measured value of offence might be the maximum) and harm by execution. However, I want to reconfirm the essential point. What specifically is the harm of homicide? Whom are we talking about when we discuss the harm of homicide? As I previously argued, citing Epicurus and his HTD, there is a metaphysical doubt about whether we should regard death as harmful. If a person simply disappears when he or she dies and death is completely harmless as HTD claims, then it seems that the retributive justification for the death penalty in terms of proportionality must be nonsense, for nothing at all happens that should trigger the process of crime and punishment. Of course, following HTD, the execution should be similarly regarded as nonsensical. However, if that is the case, the entire institutional procedure, from the perpetrator's arrest to his or her execution, must be considered a tremendous waste of time, labour, and money.

6. The notion of parsimony was newly offered to avoid a fundamental drawback of the standard retributive system, whether based on cardinal or ordinal proportionality: the standard system tends to inflict excessive, cruel punishment, as its criterion of measuring wrongness is not exempt from being arbitrary. In contrast, the newly offered system could hold inflicted punishment 'as minimally as possible, consistent with the vague limits of cardinal desert' (Walen 2015) in terms of introducing an idea of parsimony. The notion of parsimony could make the retributive system of punishment more reasonable and humane while retaining the idea of retribution.



Some may think that these kinds of arguments are merely empty philosophical abstractions. That may be. However, it is not the case that there is nothing plausible to be considered in these arguments. Consider the issue of euthanasia. Why do people sometimes wish to be euthanised? It is because people can be relieved of a painful situation by dying. That is to say, people wishing to be euthanised take death to be painless, i.e. harmless, in the same manner as HTD. This idea embedded in the case of euthanasia is so understandable that the issue of euthanasia is one of the most popular topics in ethics; however, if so, Epicurus's HTD should not be taken as nonsensical, for HTD holds in the same way as the idea embedded in the case of euthanasia that when we die, we have neither pain nor any other feeling. What I intend to highlight here is that we must be acutely aware that there is a fundamental problem concerning the notion of harm by homicide, if we want to be philosophically sincere and consistent<sup>7</sup>.

In other words, I assert that the contemporary debate over the death penalty tends to lack proper consideration for the Harm Stage in which victims themselves essentially matter, although that stage must be the very starting point of all issues. We must understand this pivotal role of the Harm Stage before intelligently discussing the death penalty. Of course, in practice, we can discuss the death penalty in a significant and refined manner without investigating the Harm Stage. For example, according to Goldman, one of the plausible positions regarding the justification for punishment in general is a position that combines both retributivism and utilitarianism. Mentioning John Rawls and H. L. A. Hart, Goldman writes (1995, p. 31):

*Some philosophers have thought that objections to these two theories of punishment could be overcome by making both retributive and utilitarian criteria necessary for the justification of punishment. Utilitarian criteria could be used to justify the institution, and retributive to justify specific acts within it.*

Goldman argues, however, that this mixed position could result in a paradox

7. Roger Crisp kindly pointed out that it is worth considering an institutional justification according to which punishment wouldn't have to be tailored to a particular case. In this view, it is sufficient that death is *generally* bad for both victims and perpetrators. I do not deny the practical persuasiveness of this view. However, from a more philosophical point of view, we should propose a question 'how can we know that death is generally bad for victims of homicide?' Following HTD, which is certainly one possible philosophical view, death is not bad at all, regardless of whether we talk about general issues or particular cases, as an agent to whom something is bad or not disappears by dying by definition. Of course, as long as we exclusively focus upon harm which the bereaved family or the society in general suffer, the institutional justification could make good sense, although in that case the issue of direct victims killed would remain untouched.

regarding how severe the punishment to be imposed on the guilty should be, even though this position avoids punishing the innocent (*ibid.*, p.36):

*While the mixed theory can avoid punishment of the innocent, it is doubtful that it can avoid excessive punishment of the guilty if it is to have sufficient effect to make the social cost worthwhile.*

This argument is useful in providing a moral and legal warning to society not to punish offenders more severely than they deserve, even if that punishment is more effective in deterring future crimes. I frankly admit that Goldman's suggestion goes to the essence of the concept of justice. However, I must also say that if his argument is applied to the death penalty, then it has not yet touched the fundamental question that forms the basis of the whole issue: whose harm should we discuss? Is it appropriate not to discuss the Harm Stage? Alternatively, I am raising the following question: who is the victim of homicide? At the very least, I think we should admit that this very question is the crucial one constituting the first problem on the death penalty, the Uncertainty of Harm.

## 6. FEELING OF BEING VICTIMISED

Next, I will examine another kind of uncertainty that is specific to the Blame Stage; the idea of retribution matters here. As far as the Japanese context for the death penalty is concerned, according to statistical surveys of public opinion, people tend to strongly support the death penalty in the case of particularly violent homicides in which they are probably feeling particularly victimised. If the death penalty were abolished, it seems that the abolition would be extremely unfair to victims of homicide, as the rights of victims (i.e. rights of life, liberty, property, and so on) would be denied by being killed, whereas those of perpetrators would be excessively protected. Obviously, the notion of retributive proportionality or equilibrium is the basis for this argument. To put it another way, this logic of retribution aims at justifying the death penalty in terms of its achieving equilibrium between the violated rights of victims and the deprived rights of perpetrators in the name of punishment. Is this logic perfectly acceptable? Emotionally speaking, I want to say yes. We Japanese might even say that perpetrators should gallantly and bravely kill themselves to take responsibility for their actions, as we have a history of the *samurai* who were expect-

ed to conduct *hara-kiri* when they did something shameful. However, theoretically speaking, we cannot accept this logic immediately, because there are too many doubtful points. Those doubts as a whole constitute the second problem concerning the death penalty.

First, we must ask, as well as in the previous section, on the issue of feeling victimised, whom are we discussing? Whose feelings and whose rights matter? Direct victims in the case of homicide do not exist by definition. Then a question arises: why can substitutes (prosecutors and others) or the bereaved family ask for the death penalty based on their feelings rather than the direct victim's feeling? How are they qualified to ask for such a stringent punishment when they were not the ones killed? The crucial point to be noted here is that the bereaved family is not identical with the direct victim. Second, even if it is admitted that the notion of the victim's emotional harm are relevant to sentencing (and at least in the sense of emotional harm the bereaved family's suffering I would agree that this makes them certainly the principal victims even if not the direct victim), it must be asked: can we justify an institution based on a feeling? This question is a part of the traditional debate concerning the moral sense theory. We have repeatedly asked whether social institutions can be based on moral sense or human feeling, when such sense or feeling cannot help but be arbitrary because those, after all, are subjective. The question is still unanswered. Third, if the feelings of being victimised justify the death penalty, then could an accidental killing or involuntary manslaughter be included in crimes that deserve the death penalty? Actually, the feelings of the bereaved family in the case of accidental killing could be qualitatively the same as in the case of voluntary homicide. However, even countries which adopt the death penalty do not usually prescribe that execution is warranted for accidental killing. Fourth, I wonder whether the bereaved family who feel victimised always desire the execution of the killer. It could be that they consider resuming their daily lives more important than advocating the execution of the murderer who killed their family member. As a matter of practical fact, executions of perpetrators need have nothing to do with supporting bereaved families. Fifth, if we accept the logic in which the death penalty is justified by the bereaved family's feeling of being victimised, how should we deal with cases where the person who was killed was alone in the world, with no family? If there is no bereaved family, then no one feels victimised. Is the death penalty unwarranted in this case? In any case, as these questions suggest, we should be aware that retributive justification based upon the feeling of being victimised is not as acceptable as we initially expected. Once again,

there is uncertainty here. Uncertainty of blame leads to the second problem concerning the death penalty.

## 7. VIOLATION AND FORFEITURE

Of course, the retributive justification for the death penalty does not have to depend upon the feeling of being victimised alone, even if the primitive basis for it might lie in human emotion. The theoretical terminology of human rights themselves (rather than emotional feeling based on the notion of rights) could be used as justification: if a person violates another's rights (to property, freedom, a healthy life, etc.), then that person must forfeit his or her own rights in proportion to the violated rights. This can be regarded as a formulation of the system of punishment established in the modern era that is theoretically based upon the social contract theory. The next remark of Goldman confirms this point (1995, p.33):

*If we are asked which rights are forfeited in violating the rights of others, it is plausible to answer just those rights that one violates (or an equivalent set). One continues to enjoy rights only as long as one respects those rights in others: violation constitutes forfeiture . . . Since deprivation of those particular rights violated is often impracticable, we are justified in depriving a wrongdoer of some equivalent set, or in inflicting harm equivalent to that which would be suffered in losing those same rights.*

However, the situation is not so simple, particularly in connection with the death penalty. In order to clarify this point, we have to reflect, albeit briefly, on how the concept of human rights has been historically established. I will trace the origin of the concept of human rights by referring to Fagan's overall explanation. According to Fagan (2016, Section 2):

*Human rights rest upon moral universalism and the belief in the existence of a truly universal moral community comprising all human beings . . . The origins of moral universalism within Europe are typically associated with the writings of Aristotle and the Stoics.*

Followed by the remark:

*Aristotle unambiguously expounds an argument in support of the existence of a natural moral order. This natural order ought to provide the basis for all truly rational systems of justice . . . The Stoics thereby posited the existence of a universal moral community effected through our shared relationship with god. The belief in the existence of a universal moral community was maintained in Europe by Christianity over the ensuing centuries.*

This classical idea was linked during the 17th and 18th centuries to the concept of ‘natural law’ including the notion of ‘natural rights’ that each human being possesses independently of society or policy. ‘The quintessential exponent of this position was John Locke . . . Locke argued that natural rights flowed from natural law. Natural law originated from God’ (ibid.). Fagan continues (ibid.):

*Analyses of the historical predecessors of the contemporary theory of human rights typically accord a high degree of importance to Locke’s contribution. Certainly, Locke provided the precedent of establishing legitimate political authority upon a rights foundation. This is an undeniably essential component of human rights.*

Although, of course, we should take post-Lockean improvement including Kantian ideas into account to fully understand contemporary concepts of human rights, we cannot deny that Locke’s philosophy ought to be considered first.

As is well known, Locke’s argument focuses on property rights. He put forth the idea that property rights were based on our labour. Thus, his theory is called ‘the labour theory of property rights’. Let me quote the famous passage I have in mind (Locke 1960, Second Treatise, Section 27):

*Though the Earth, and all inferior Creatures be common to all Men, yet every Man has a Property in his own Person. This no Body has any Right to but himself. The Labour of his Body, and the Works of his Hands, we may say, are properly his.*

This idea could cover any kind of human rights such as those for living a healthy life, liberty, and property, because human rights are supposed to be *owned* by us. For example, H. L.A. Hart once argued that legal rights are nothing but legal powers to require others to meet correlative obligations, and then pointed out that; ‘we also speak of the person who has the correlative right as *possessing* it or even *owning* it’

(Hart 1982, p.185). If this is the case, we can make property rights representative of all human rights.

However, if we follow Locke's theory (and many countries, including Japan, still do), then it logically follows that what we cannot gain by our labour by definition cannot be objects of human rights. How does Locke's idea apply to our life itself (rather than simply living a healthy life)? Are we able to acquire our life itself by our labour? No, we cannot. We can realise a healthy life by making an effort to be moderate, but we cannot create our lives. We are creatures or animals; therefore, our lives are not something that we ourselves made by our labour. Locke uses the concept of power (as Hart does) when he discusses various aspects of property rights. Among those, we should pay particular attention to the following (Locke 1960, Second Treatise, Section 23):

*For a Man, not having the Power of his own life, cannot, by Compact, or his own Consent, enslave himself to any one, nor put himself under the Absolute, Arbitrary Power of another, to take away his Life, when he pleases.*

Locke also writes (1960, Section 24):

*No Man can, by agreement, pass over to another that which he hath not in himself, a Power over his own life.*

Obviously, Locke assumes that we have no property rights over our own lives or bodies themselves, or more precisely, no property rights in controlling and destroying our own lives as a whole; therefore, we cannot alienate those rights to others. We cannot alienate or forfeit what we do not have. If this is the case and we presuppose the formulation of the system of punishment introduced above in terms of violation and forfeiture, what would result? The answer is clear. Our lives themselves are conceptually beyond the terminology of human rights, and thus, if the death penalty is defined as a punishment requiring the forfeiture of the perpetrator's right to life, the death penalty should be regarded as conceptually contradictory or impossible. We cannot lose tails, as we do not have tails. Likewise, we cannot own our lives (i.e. we have no property rights in our life itself), so we cannot lose our lives, at least in such a sense as forfeiture of human rights. This is the third route to an 'impossibilist' view of the death penalty. This argument depends heavily on Locke's original theory.



Nevertheless, as long as we have to consider Locke's classical view seriously in order to discuss the relation between punishment and human rights, we must be aware that we could be involved in theoretical uncertainty in justifying the death penalty through the notion of human rights in a retributivist flavour, as the argument thus far suggests. This is the very puzzle that I want to propose as the third problem concerning the death penalty debates.

Moreover, we must acknowledge that retributive ideas in the Blame Stage usually include a kind of evaluation of the psychological state of the agent's behaviour at the time of the crime as a matter of legal fact. In other words, rationality, freedom, or *mens rea* are usually needed for agents to be judged guilty. However, from a strictly philosophical perspective, we should say that it is far from easy in principle to confirm those states in the past. Indeed, this psychological trend seems to cause controversy in court proceedings, as seen, for example, in the American context known as 'battered-woman syndrome'. If a woman who has been routinely battered by her partner suddenly fights back and kills her partner, American courts often find her not guilty. People wonder whether such an evaluation concerning battered women could be correctly made without arbitrariness. Additionally, philosophical debates on free will and the development of the brain sciences must be considered. Some philosophers assert that we have no free will because our personality and actions are intrinsically governed by external factors, such as our environments or biological conditions, which are definitely beyond our control. This philosophical standpoint is often called 'hard incompatibilism' (see Strawson 2008). In this respect, my analogy to a natural disaster could be seen as appropriate, as our actions might be taken to be just natural phenomena at the end of the day.<sup>8</sup> Furthermore, brain sciences often provide shocking data to suggest that our will may be controlled by brain phenomena occurring prior to our consciousness, as shown by Benjamin Libet. In view of such contemporary arguments, we have little choice but to say that we cannot be perfectly certain whether a given perpetrator who committed homicide is truly guilty, as long as we adopt the present standard for judging the psychological states of offenders in court. To sum up, the third problem for the death penalty is the difficulty in knowing

8. Additionally, my analogy with natural disasters, particularly the case of the 2011 quake, could be re-confirmed to be appropriate in the sense of presenting a similar kind of uncertainty to the case of the death penalty. The danger of constant exposure to low doses of radiation for long periods involves some uncertainty, as far as we now know. Fortunately, however, the dose of radiation to which the people of Fukushima were exposed as a result of the 2011 quake, internally and externally, was low enough for us to be certain, based upon past epidemiological research, that no health problems will arise in the future. Regarding radiation exposure, everything depends upon the level of dose. The smaller the dose, the less dangerous it is.

whether someone has property in their life itself as well as uncertainty about the mental state of the accused, this is the Uncertainty of Rights Violation.

## 8. THE DETERRENT EFFECT

Finally, I will examine some problems in the Danger Stage. What matters in this context is the utilitarian justification for the death penalty; I will focus on what is called the ‘deterrent effect’. Firstly, I would like to say that the death penalty undoubtedly has *some* deterrent effect. This is obvious if we imagine a society where violators of any laws, including minor infractions such as a parking ticket or public urination, must be sentenced to death. I believe that the number of all crimes would dramatically reduce in that society, although it would constitute a horrible dystopia. The argument for the deterrent effect of the death penalty probably arises from the same line of ‘common sense’ thinking. For example, Pojman says, ‘there is some non-statistical evidence based on common sense that gives credence to the hypothesis that the threat of the death penalty deters and that it does so better than long prison sentences’ (Pojman 1998, pp. 38-39). Specifically, this deterrent effect presupposes the utility calculus that a human being conducts, whether consciously or unconsciously, in terms of ‘weighing the subjective severity of perceived censure and the subjective probability of perceived censure against the magnitude of the desire to commit the offence and the subjective probability of fulfilling this desire by offending’ (Beyleveld 1979, p. 219). Therefore, if we presuppose the basic similarity of human conditions, it may be plausible to state the following about the deterrent effect of punishment: ‘this can be known *a priori* on the basis of an analysis of human action’ (ibid., p. 215). However, in fact, the death penalty in many countries is restricted to especially heinous crimes, such as consecutive homicides (although some countries apply the death penalty to a wider range of crimes), which suggests that we must conduct empirical studies, case by case, if we want to confirm the deterrent effect of the death penalty. Therefore, the question to be asked regarding the deterrent effect is not whether the death penalty is actually effective, but rather how effective it is in restricted categories of crimes. What matters is the degree.

There are many statistical surveys concerning this issue. In particular, an economic investigation by Ehrlich is often mentioned as a typical example. After examining detailed statistical data and taking into account various factors, such as race, heredity, education, and cultural patterns, Ehrlich suggests (1975, p. 414):

*An additional execution per year over the period in question [i.e., 1935-1969] may have resulted, on average, in 7 or 8 fewer murders.*

Of course, this estimate includes too many factors and presumptions to be perfectly correct. Ehrlich himself is aware of this and thus says (ibid.):

*It should be emphasized that the expected tradeoffs computed in the preceding illustration mainly serve a methodological purpose since their validity is conditional upon that of the entire set of assumptions underlying the econometric investigation ... however ... the tradeoffs between executions and murders implied by these elasticities are not negligible, especially when evaluated at relatively low levels of executions and relatively high level[s] of murder.*

Ehrlich's study drew considerable criticism, most of which pointed out deficiencies in his statistical methodology. Therefore, at this moment, we should say that we are able to infer nothing definite from Ehrlich's study, although we must value the study as pioneering work.

Van den Haag proposes an interesting argument based upon uncertainty specific to the deterrent effect of the death penalty. He assumes two cases, namely, case (1), in which the death penalty exists, and case (2), in which the death penalty does not exist. In each case there is risk or uncertainty. On the one hand, in case (1), if there is no deterrent effect, the life of a murderer is lost in vain, whereas if there is a deterrent effect, the lives of some murderers and innocent victims will be saved in the future. On the other hand, in case (2), if there is no deterrent effect, the life of a convicted murderer is saved, whereas if there is a deterrent effect, the lives of some innocent victims will be lost in the future (Van den Haag 1995, pp. 133-134). Conway and Pojman explain this argument using the following table, 'The Best Bet Argument', which I have modified slightly, having DP stand for the death penalty, and DE the deterrent effect:

## THE WAGER

	<i>DE works</i>	<i>DE does not work</i>
<i>We bet DP works</i>	save: murderers and innocent victims in the future lose: convicted murderer	save: nothing affected lose: convicted murderer
<i>We bet DP does not work</i>	save: convicted murderer lose: innocent victims in the future	save: convicted murderer lose: nothing affected

Following this table, Conway assumes (after Van den Haag's suggestion that the life of a convicted murderer is not valued more highly than that of the unknown victims) numerical values about each case (each numerical number stands for not a number of people but a hypothetical value for a person to be saved or killed) :

- a murderer saved +5
- a murderer executed -5
- an innocent saved +10
- an innocent murdered -10

Moreover, he assumes that for each execution, only two innocent lives are spared (i.e. he assumes the deterrent effect to be almost the minimum). Then, consequently, executing convicted murderers turns out to be a good bet (Conway 1995, pp. 265-266; Pojman 1998, pp. 40-41).

## 9. NEGATIVE CAUSATION AND WHERE TO GIVE PRIORITY

Van den Haag's 'Best Bet Argument' sounds quite interesting. However, Conway has already proposed a fundamental challenge to this argument: it mistakenly regards the *actual* death of convicted murderers as being on a par with the *possible* death of innocent victims in the future (Conway 1995, pp. 269-270). This is confusing or possibly a rhetorical sleight of hand. I think that Conway's reaction to Van den Haag's argument is a reasonable one.

As I approach my conclusion, I will propose two problems with Van den Haag's argument. First, I want to acknowledge that any arguments, including Van den Haag's, supporting the death penalty in terms of its deterrent effect seem to presuppose a causal relationship between the existence of the death penalty and people not killing others. For example, Pojman writes, 'the repeated announcement and regular exercise of capital punishment may have deep causal influence' (1998, p. 48). However, epistemologically speaking, that presupposition is extremely hard to confirm, because the effect of this causal relationship is not a positive, but rather a negative event, which is the event of not killing others. This has something to do with the philosophical problem of how to understand negative properties. By negative properties we mean that, for example, my room is not full of seawater; my room does not consist of paper; my room is not melting us, etc. Such descriptions by negative properties can be made almost endlessly. In other words, one identical event described by a positive property (e.g., this room is well lit) can be re-described in infinite ways in terms of negative properties. Take the example that I am now at my computer in Tokyo, writing a paper. This event can also be described as 'I am not eating', 'I am not sleeping', 'I am not killing others' (!), etc. The positive event, 'I am writing a paper now', can be understood through a causal relationship. The event was most likely caused by my intention to do so, which was caused by my sense of duty as a professor, etc. How, then, could we understand the negative description of my action, 'I am not killing others'? Was this caused by the existence of the death penalty in Japan?

Perhaps I was completely unaware of the existence of the death penalty in Japan when I wrote a paper without killing others. Could the death penalty be its cause? Could the negative event 'I am not killing others' be an effect of the death penalty? It is hard to say so.

This problem is the same as the problem of 'causation by absence' or 'omission-involving causation'. Generally, causation by absence is usually examined in the form of answering a question about whether nothingness can cause something. For example, David Lewis discusses a question about how a void (understood as being entirely empty or nothing at all, differing from a vacuum) is regarded as a cause of something (Lewis 2004). He says, 'If you were cast into a void, it would cause you to die in just a few minutes. It would suck the air from your lungs. It would boil your blood. It would drain the warmth from your body. And it would inflate enclosures in your body until they burst' (ibid., p.277). However, the problem is that the void is nothing. 'When the void sucks away the air, it does not exert an attractive force on

the air' (ibid.). Furthermore, another, perhaps harder problem would arise. We can say, 'If I defended you from being cast into a void, you would not die'. Namely, my omission to defend you would cause you to die. However, should only my omission matter? What of your brother's omission to defend you? Or the Prime Minister of the UK's omission to defend you? Are not all of those qualified to be the cause of your death, as least as long as we adopt a common-sense counterfactual analysis of causation? As this argument suggests, in the context of the current debate on this problem, the most troublesome phase is that 'too many' absences can be supposed to cause a particular effect. I quote Menzies, who says (2004, p.145):

*I am writing this essay at my computer. If, however, there were nerve gas in the air, or I were attacked with flamethrowers, or struck by a meteor shower, I would not be writing the essay. But it is counterintuitive to say that the absence of nerve gas, flamethrower attack, and meteor strike are causes of my writing the essay.*

This example takes the issue of absence as a cause, but simultaneously his example refers to the case of effect as absence (not writing the essay). As this shows, the current debate on the problem of causation by absence could extend to the case of effect as absence. In any case, what matters is a possibility that 'too many' absences can cause something, and something can cause 'too many' absences (Menzies calls this problem 'the problem of profligate causation' (ibid., pp.142-145)). Then the deterrent effect of the death penalty is definitely classified as a case of absence as effect rather than cause. In other words, the absence of homicide (as effect) matters, whereas in this case execution (as cause) is presupposed to exist. It seems that the current debate on causation by absence is highly likely to contribute to discussing the problem of the deterrent effect.

Of course, someone may counter my argument by saying that what matters in this context is a statistical correlation between the number of executions and the number of homicides, which could be confirmed in an empirical way. I admit that the statistical correlation plays a crucial role here, even though we must simultaneously acknowledge that what is called 'randomized controlled trial', the most reliable, statistical methodology to confirm causal relations, is unfeasible due to the nature of the problem. Actually, this kind of correlation is too rough to predict the causal relationship between those, although the causation really matters. Causes of a reduction or increase in the number of homicides can be interpreted or estimated in various ways,



considering confounding factors, such as education, economic situation, urban planning, and so on. Therefore, in principle, there always remains the possibility that the apparent correlation between the death penalty and the reduction of homicides is merely accidental. For example, there may be another, common cause, that brings about both people's tendency to support the death penalty and the reduction of homicides<sup>9</sup>. We should recognise that there is intrinsic uncertainty here. These difficulties concerning causal relations give rise to a fourth problem related to the death penalty debates – the Uncertainty of Causal Consequences.

Incidentally, let me now return to my distinction of the three stages regarding the death penalty. Obviously, the issue of the deterrent effect belongs primarily to the Danger Stage. Yet it is vital to consider the Harm Stage. How can the deterrent effect affect the Harm Stage? I must say that the retentionist's argument, in terms of the deterrent effect of the death penalty, completely dismisses this essential point. We need only recall the analogy of the 2011 quake in Japan. 'Retentionism' based upon the deterrent effect corresponds to aspect (3), where the improvement of the preventive system matters. This is important, of course, but cannot be a priority. Priority lies in the issues of how to deal with the actual harm that the victims have already suffered (specifically referring to the bereaved family or others in the case of homicide and the death penalty). Without consideration of how to cope with the harm, even if the theory seriously considers the innocent victims in the future, the retentionists' theory can hardly be persuasive.

It is true that the retentionists' theory based on the deterrent effect appropriately considers the person harmed in the process of punishment. For example, Walker considers such a phase in the process of punishment as one of the possible objections against retentionism based on the deterrent effect by saying: 'if the benefit excludes the person harmed this too is nowadays regarded by many people as morally unacceptable' (Walker 1980, p. 65). However, as the context clearly shows, by 'the person harmed' he means the person punished. He does not mention the initial harm suffered by victims. This problem is concerned with my previous claim; that is, we have to consider the 'whom-question' when we discuss the justification of punishment. Whom are we discussing? Whose benefit do we consider? In the face of victims before our eyes, can we emphasise only the improvement of preventive systems for

9. On negative causation and the possibility of common cause, see Ichinose (2013). In particular, my argument on negative causation concerning the death penalty rests on my argument of Ichinose (2013).

the future? Evidently, actual victims are the first to be helped, although obviously it is not at all bad to simultaneously consider the preventive system in the future. It is necessary for us to respect basic human rights and the human dignity of perpetrators and innocent people in the future; however, that respect must be in conjunction with our first taking care of actual victims. We ought not to get our priorities wrong.

## 10. PROSPECTS

I have indicated that the debates on the death penalty are inevitably surrounded by four problems over specific kinds of uncertainties: uncertainty concerning the victim of homicide, uncertainty in justifying the death penalty from the feeling of being victimised, uncertainty in justifying the death penalty on the basis of human rights, and uncertainty over negative causation. In the course of examining these problems, I have proposed the option of developing an ‘impossibilist’ position about the death penalty, which I am convinced, deserves further investigation. However, being surrounded by theoretical problems and uncertainties might be more or less true of any social institution. My aim is only to suggest how the death penalty should be understood as involving uncertainties from a philosophical perspective. Most likely, if there is something practical that I can suggest based on my argument, then what we might call a ‘Harm-Centred System’ may be introduced as a relatively promising option instead of, or in tandem with, the death penalty. What I mean by this is a system in which we establish as a priority redressing actual harm with regard to legal justice, where ‘actual harm’ only implies what the bereaved family suffer from, as the direct victims have already disappeared in the case of homicide. In other words, I think that something akin to the maximalist approach to restorative justice<sup>10</sup> or some hybrid of the traditional justice system and the restorative justice system should be seriously considered, although we cannot expect perfect solutions exempt from all of

10. According to Bazemore and Walgrave, ‘restorative justice is every action that is primarily oriented towards doing justice by repairing the harm that has been caused by a crime (Bazemore and Walgrave 1999 (2), p.48). Restorative justice, that is to say, is a justice system that mainly aims at restoring or repairing the harm of offences rather than punishing offenders as the retributive justice system does. Initially, restorative justice has been carried out by holding ‘a *face-to-face* meeting between the parties with a stake in the particular offense’ (ibid.) like victim, offenders, or victimised communities. However, this type of justice system works only in a complementary way to the traditional system of retributive justice. Then, the maximalist approach to restorative justice was proposed, which seeks to develop ‘restorative justice as a fully-fledged alternative’ (Bazemore and Walgrave 1999 (1). Introduction. P.8) to retributive justice. This approach ‘will need to include the use of coercion and a formalization of both procedures and the relationship between communities and society’ (ibid., p.9.)

the above four problems. It is certainly worth considering whether some element of restorative justice can play a significant role in the best theory of punishment.

In any case, my argument is at most a philosophical attempt to address problems. How to apply it to the practice of the legal system is a question to be tackled in a future project.

#### REFERENCES

- Bazemore, G. and Walgrave, L. 1999 (1). 'Introduction: Restorative Justice and the International Juvenile Justice Crisis'. In *Restorative Juvenile Justice: Repairing the Harm of Youth Crime*, eds. G. Bazemore and L. Walgrave, Criminal Justice Press, 1-13.
- . 1999 (2). 'Restorative Juvenile Justice: In Search of Fundamentals and an Outline for System Reform'. In *Restorative Juvenile Justice: Repairing the Harm of Youth Crime*, eds. G. Bazemore and L. Walgrave, Criminal Justice Press, 45-74.
- Beyleveld, D. 1979. 'Identifying, Explaining and Predicting Deterrence'. *British Journal of Criminology* 19:3, 205-224.
- Calvert, B. 1993. 'Locke on Punishment and the Death Penalty'. *Philosophy* 68:264,, 211-229.
- Collins, J., N. Hall, and L. A. Paul. 2004. *Causation and Counterfactuals*. MIT Press.
- Conway, D. A. 1995 (originally 1974). 'Capital Punishment and Deterrence: Some Considerations in Dialogue Form'. In *Punishment: A Philosophy and Public Affairs Reader*, eds. J. Simmons, M. Cohen, J. Cohen, and C. R. Beitz. Princeton University Press, 261-273.
- Diogenes Laertius. 1925. *Lives of Eminent Philosophers*. Vol. 2. Trans. R. D. Hicks. Loeb Classical Library. William Heinemann Ltd.
- Ehrlich, I. 1975. 'The Deterrent Effect of Capital Punishment: A Question of Life and Death'. *American Economic Review* 65:3, 397-417.
- Fagan, A. 2016. 'Human Rights'. In Chase B. Wrenn, *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002. Available from <http://www.iep.utm.edu/hum-rts/#H2> [Accessed 12 June 2017]
- Fischer, J. M., ed. 1993. *The Metaphysics of Death*. Stanford University Press.
- Goldman, A. H. 1995 (originally 1979). 'The Paradox of Punishment'. In *Punishment: A Philosophy and Public Affairs Reader*, eds. J. Simmons, M. Cohen, J. Cohen, and C. R. Beitz. Princeton University Press, 30-46.
- Hart, H. L. A. 1982. *Essays on Bentham: Jurisprudence and Political Theory*. Oxford University Press.

Ichinose, M. 2013. 'Hybrid Nature of Causation'. In T. Uehiro, *Ethics for the Future of Life: Proceedings of the 2012 Uehiro-Carnegie-Oxford Ethics Conference*, the Oxford Uehiro Center for Practical Ethics, University of Oxford, 60-80.

———. 2016. 'A Philosophical Inquiry into the Confusion over the Radiation Exposure Problem'. *Journal of Disaster Research* 11: No.sp, 770-779.

Lewis, D. 2004. 'Void and Object'. In J. Collins, N. Hall, and L. A. Paul, *Causation and Counterfactuals*. MIT Press, 277-290.

Locke, J. 1960. *Two Treatises of Government*, ed. P. Laslett, Cambridge University Press.

———. 1975. *An Essay concerning Human Understanding*, ed. P. H. Nidditch. Oxford University Press.

Menzies, P. 2004. 'Difference-Making in Context'. In J. Collins, N. Hall, and L. A. Paul, *Causation and Counterfactuals*. MIT Press, 139-180.

Pojman, L. P., and J. Reiman. 1998. *The Death Penalty: For and Against*. Rowman & Littlefield Publishers, Inc.

Scanlon, T. M. 1982. 'Contractualism and utilitarianism'. In A. Sen and B. Williams, *Utilitarianism and Beyond*. Cambridge University Press, 103-128.

Simmons, A. J. 1994. 'Locke on the Death Penalty'. *Philosophy* 69:270, 471-477.

Strawson, G. 2008. 'The Impossibility of Ultimate Moral Responsibility'. In *Real Materialism*. Oxford University Press, 319-331.

Tonry, M. 1994. 'Proportionality, Parsimony, and Interchangeability of Punishments'. In *A Reader on Punishment*, eds. A. Duff and D. Garland. Oxford University Press, 133-160.

Van den Faag, E. 1995 (originally 1969). 'On Deterrence and the Death Penalty'. In *Punishment and the Death Penalty: The Current Debate*, eds. R. M. Baird and S. E. Rosenbaum. Prometheus Books.

Von Hirsch, A. 1993. *Censure and Sanctions*. Oxford University Press.

Walen, A. 2015. 'Retributive Justice'. *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.). Available from <http://plato.stanford.edu/archives/sum2015/entries/justice-retributive/> [Accessed 12 June 2017]

Walker, N. 1980. *Punishment, Danger and Stigma: The Morality of Criminal Justice*. Barnes & Noble Books.