University of Reading

Geoengineering Governance: Addressing the Problems of Moral Corruption, Moral Hazard, and Intergenerational Inclusion.

PhD thesis in Political Theory

School of Politics Economics and International Relations

Joshua Wells

January 2020

Declaration of Authorship

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Signed: Joshua Wells

Date: 06/01/2020

Acknowledgements

‘I am not going to thank anyone because I did it all by myself’ (Spike Milligan).

As tempting as it is for the above to be the extent of my acknowledgements, it would be far from true. I consider myself to be lucky to have the positive impact of many people in my thesis. Here I will name some of whom I feel particularly grateful.

Firstly I would like to acknowledge my gratitude to my supervisors, whose fingerprints exist on every page of the thesis. My primary supervisor Catriona McKinnon, inspired my interest in climate justice as a second year undergraduate in her political thinking module. During the supervising of my thesis she has provided perpetual guidance and insights as well as creating an amazing environment to do the PhD with the Leverhulme climate justice program. My second supervisor, Alice Baderin has been brilliant since joining my project in November 2018; she has provided a huge amount of constructive input to the thesis; I failed to appreciate the value of a second supervisor until she took up the role.

The Leverhulme Climate Justice program has provided a remarkable environment in which to undertake my PhD. I am grateful to the Leverhulme trust for giving me the opportunity to partake in this program. One of the key benefits being the other scholars on the program some of whom have provided consistently helpful comments, such as Adam Pearce, Bennat Francis, Daniel Harris, Jamie Draper, Livia Luzzatto, Lydia Messling and Vera Van Gool.

The following people deserve special thanks, for they have endured a disproportionate share of my thinking out loud to them about the thesis, as well as reading and enduring some of my written work, and they have given invaluable helpful advice for the completion of the thesis. They are Aart Van Gils, Alex McLaughlin, Christos Vrakopoulos, Michela Bariselli, and Merve Altintas. The thesis would be significantly different without them.

Reading is a brilliant community, in more general terms I have benefited from chatting with, and input from: Brian Feltham, Faye Bird, Jon Elsey, Julia Mosquera, Julius Kapembwa, Marat Shardimgaliev, Marco Bernardini, Marina Della Giusta, Matthew Robson, Neha Hui, Rob Jubb, Sarah Najm, Tom Lee and Yanos Soubieski.

At Reading I consider myself to be very lucky to have amazing friends who have achieved the feat of making the experience of doing a PhD enjoyable, I would like to acknowledge a few: Adnan Farock, Alice Romano, Amin Fatemi, Antonio Scarafone, James Wagstaff, Laura

Bucher, Liliana kamaka, Luisa Campi, Martina Mutti, Natalia Kampakli, Nico Biagi, Nick Byrne, Nuri Altintas, Rebecca Rusk Steve Smith, Zihad Khan, and Zoi Vardanika.

Reading also has a great graduate school, which has been a great place to work on this thesis. I am especially grateful to Helen Apted and Nicky Stepney for making the graduate school such a pleasant place to work. The final feature of Reading to mention are those who have made it such a good place by being willing to play sports (particularly squash), I cannot list the people I have played but they have made my time here significantly happier.

I also consider myself fortunate that my thesis was examined by Alan Cromartie and Clare Heyward, who have challenged me to think more carefully about the thesis, and the thesis now reflects the benefit of their careful and constructive scrutiny.

Beyond the confines of Reading, I have been lucky enough to meet talent political theorists with whom I have had illuminating conversations, and received valuable feedback from presenting to them. To acknowledge a few particularly helpful ones I am grateful to the advice of: Anna Wienhues, Ben Hale, Clare Heyward, Daniel Callies, Dominic Roser, Eike Düvel, Harold Stezler, Henry Shue, John Barry Kian Mintz-Woo, Simon Caney, and Stephen Gardiner.

I would also like to acknowledge the role of some people who challenged me when I started getting interested in Politics, and by doing so showed me the joy of political conversation, without this I doubt I would have had the interest in the subject to study it. These people are Andrew Shadrake, David Horne, Geoff Campbell-Follet and Phil Chappell. In a similar vein I appreciate the role of some of my friends outside Reading who have been supportive, encouraging and tolerant of me developing my political interests, particularly, Edward Hopkins (and the whole Hopkins family), Matthew Hill, Simon Farnsworth, Richard Smith, Robin Addison and William Honeywell.

Additional there are those who encouraged and me to start the PhD, Fay Farstad and Paul Tobin who were remarkably supportive in this regard; without them I doubt I would have applied. Moreover there are Phil and Ros Johnson who were remarkably generous with their support when I started the thesis.

Finally I would like to thank my family, my granddad (Jack) showed the importance of trying to understand a subject in more depth and the fulfilment which comes from this. The other Jack in my family, my dog provided a source of joy and support regardless of how the thesis was

going, and reminded me of the importance of having fun as well as thinking about non-human animals. Additionally, this is only possible with the perpetual support and tolerance of my brother Tom, my father Bill, and my mother Lynne, for supporting me throughout the thesis which must have been comparable to watching a slow version of the never ending story.

All the normal disclaimers apply to these acknowledgements, I would like to claim the credit for the errors and blame others for the merits of the thesis.

Abstract

This thesis is concerned with how geoengineering, specifically Stratospheric Sulphur aerosol Injection (SSI), could be ethically governed. Geoengineering refers to a set of technologies which can be used to affect the global temperature, such as SSI. Geoengineering has been proposed as possible response to climate change. This thesis focuses on three problems which the ethical governance of SSI faces: namely the problems of moral corruption, moral hazard and intergenerational inclusion. The result of this is that the thesis furthers our understanding of how SSI governance could address each of these three problems. By doing so the thesis contributes t an important debate on how SSI should be governed.

The first chapter presents a case in favour of the importance of the ethical governance of SSI. Chapter two introduces two framings which are seldom used together, the risk-risk trade off frame, and the perfect moral storm frame. Chapter two argues that it is important to adopt both of these frames if we are going to consider geoengineering governance. The benefit of these frames is that they provide a context against which thinking about SSI governance occurs.

Chapter three explains the problem of moral corruption, and argues that a well-functioning accountability mechanism could help to address it. In making this argument, the chapter shows that transparency, publicity, and accountability are poorly understood in reports on geoengineering governance, in which these principles are often endorsed. The chapter offers a clearer account of the meaning of these principles, and why it is essential to be aware of the relationship between transparency and these other principles if they going to be used to address moral corruption or any problem in SSI governance.

Chapter four provides conceptual clarity to the often-cited moral hazard concern about SSI. The chapter breaks the moral hazard problem down into five different variables. By doing so this analysis highlights the lack ambiguity and disagreement in the literature about the moral hazard.. This chapter also provides an answer to the question of should we act on the hazard if the empirical evidence about the hazard effect is inconclusive. Drawing on the work of Henry Shue on threshold likelihoods, I argue that we should act on the moral hazard problem even if the empirical evidence is inconclusive, due to the mechanisms by which the hazard can occur being well-understood, and that these mechanism are accumulating. .

Chapter five explores the possibility of secrecy as a response to the moral hazard concern. The chapter has four components. Firstly it provides clarity about how secrecy can be understood. Secondly, it considers what this theoretical account of secrecy teaches us about the practice of

secrecy. This is done by applying the theory to the historical example of the Manhattan project, which highlights some key components of governing in secrecy. Thirdly, the chapter shows us why we should expect such an approach to secrecy to be effective at addressing the moral hazard problem. Fourthly, the chapter has a normative component, whereby it considers instrumental reasons to be opposed to secrecy in SSI governance despite its promise in addressing the moral hazard problem.

Chapter 6 considers the question of intergenerational inclusion in decision-making about SSI. It identifies representation of interests as the appropriate form of inclusion for future people. It argues in favour of '*A Statement of What is Owed to the Future*', whereby the minimal interests of future generations are expressed and accepted by states. The chapter proceeds by considering different mechanisms through which these interests could be represented, and argues in favour of a second chamber as the ideal mechanism through which this can occur.

. The idea of ethical SSI governance is complicated and confusing, it faces many challenges, and we face a genuine risk of SSI being governed in a poor or unethical way. Even if agents wished to govern SSI ethically it is not at all clear how this can be done. This thesis makes the prospect of ethical SSI governance more attainable by helping agents understand what to do in light of three important ethical problems SSI governance ought to address.

Contents

1 The Context: Why Consider Geoengineering?

It is hard to overemphasise how important the climate is. An obvious statement perhaps, but one I plan to expand on and explore. The climate is a background condition of life. Amongst other things, it can determine whether we have access to water and food, it can affect what type of shelter we require, and it is an indicator of what type of weather we will experience.[1] Insofar as we value life, we clearly have an interest in the climate. The study of interplanetary meteorology shows this concern. (Horizon, The Wildest Weather in the Universe: 2016). One of the first questions to be asked regarding newly discovered planets is: what is the climate there like? The reason why this question is asked is it helps to determine what the possibility of life is, and what types of life could exist on that planet.

The Earth's climate is changing. This means that the background condition of life is changing. The climate is changing faster than we can adapt to the changes and this represents one of the greatest threats of the twenty-first century to human and non-human life. Climate change will affect our lives, the lives of the vulnerable, the lives of those who are yet to be born, as well as present and future non-human life and the natural world. We are fortunate, however, because we know how this change is happening. Since 1990 policy-makers from across the world have acknowledged that humans are the driving force behind this change, due to our emissions of greenhouse gases (Jamieson, 2014: 33).

As humans alive today, we know that our emissions represent a substantial part of the problem since they are contributing to climate change. Furthermore, this puts humanity in an incredible position: the twenty-first-century threat to human survival is one of our own making. Moreover, it is arguably one that we can stop. We can choose to put a stop to carbon emissions, and the threat of climate change would be reduced. Unfortunately, this would not prevent the changes to which are already committed, as greenhouse gases have long lives. Consider the carbon dioxide emitted from fossil fuels: 20–60% of carbon dioxide remains in the atmosphere for 1,000 years or longer (Archer and Brovkin, 2008). This means that even if net global carbon emissions were to become zero today and stay at zero, there would still be some form of climate change due to the life cycle of the carbon which is already in the atmosphere.[2] This changes

[1] This concern with weather is not the British concern about whether it will be hot enough to wear shorts. It is (in its most extreme form) a concern about how weather conditions threaten lives.

[2] This is why the language around mitigation and adaptation dominates the literature on responses to climate change, as opposed to prevention.

the scale of the problem entirely. Moreover, we are fortunate enough to know how to minimise the risks we are placing on one of our background conditions of life.

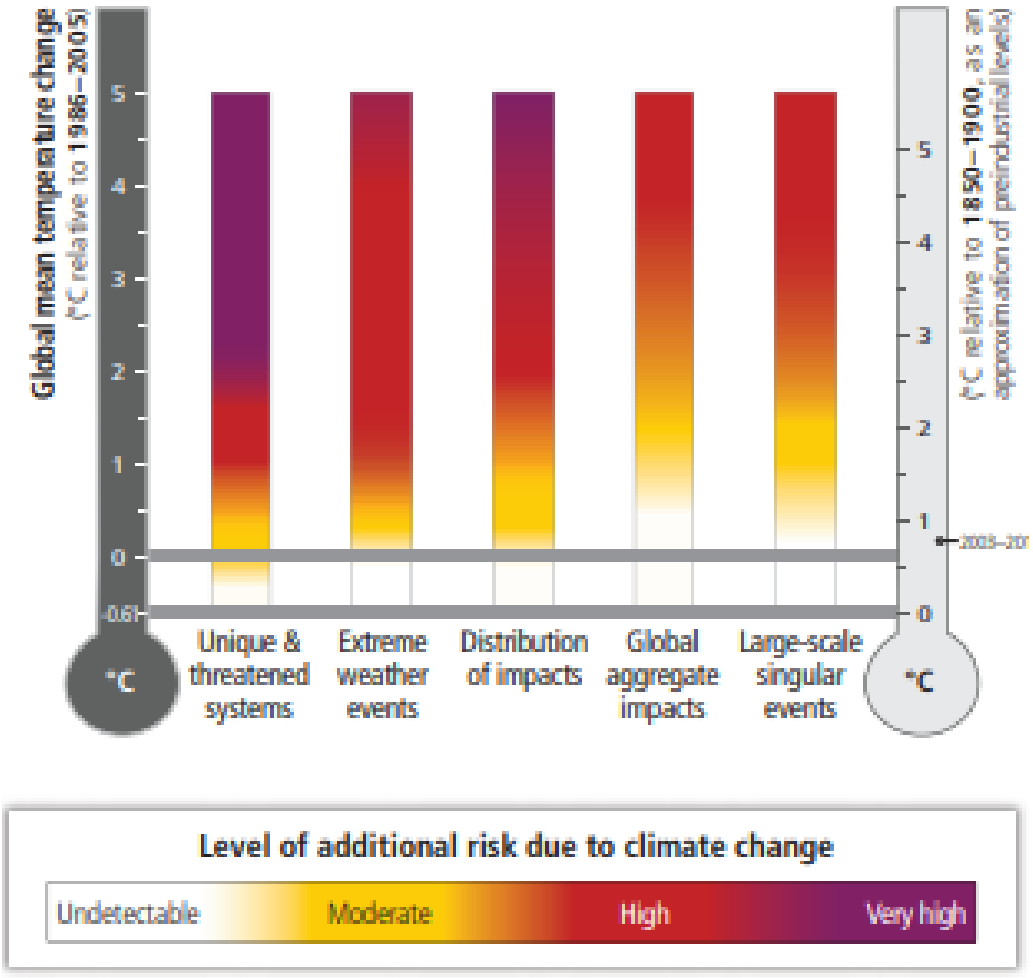
Unfortunately, the situation is not so simple. Despite policy-makers being aware of the threat of climate change since the early 1990s, action on the problem has not been enough to ensure that the change does not threaten the background conditions of life for all. The conversation in policy circles has focused on the idea of ensuring that there is no more than two degrees Celsius (2^oC) of warming (Carbon Brief, 2014). Regrettably, it is doubtful that the limit of 2 degrees will be met, given that our current emissions trajectory appears to commit the world to warming of between 2.6^oC and 3.1^oC (Rogelj *et al*., 2016: 634).[3]

The Working Group II contribution to the Intergovernmental Panel on Climate Change (IPCC) in 2014 provides a framework for understanding what is bad about this level of warming. Specifically, the Working Group II identifies 5 'reasons for concern' about projected levels of climate change. (1) The danger is that ecosystems and cultures, some of them unique, are being threatened by climate change and that the number and intensity of these threats are increasing as the global mean temperature rises. (2) Extreme weather events such as heatwaves and coastal flooding are highly likely to occur more often and with a greater degree of intensity as global temperatures rise. (3) The risks from climate change and its impact are unevenly distributed and likely to have a greater impact on vulnerable communities. This is in part due to regional variations in the impact of climate change on important necessities of life such as the availability of water. (4) Global aggregate impacts relate to the concern that global phenomena such as biodiversity and the global economy will suffer. Climate change is already causing a loss of biodiversity and having an impact on the global economy, and these impacts are likely to grow as the global mean temperature rises. (5) The concern with large-scale singular events refers to sudden and irreversible changes to the Earth's physical or ecological systems such as the melting of the Greenland ice sheet, which would cause up to a 7-metre increase in the global mean sea level.

The IPCC Working Group II provides the graph in Figure 1, which shows how these five 'reasons for concern' compound in likelihood as the global mean temperature increases.

[3] Due to a commitment from states for more ambitious nationally-determined contribution levels, this number could become lower. However, it is also premised on states acting on their emissions reduction pledges agreed in Paris 2016, which many have yet to do. Hence the figure may also be significantly higher.

Figure 1. Five Reasons for Concern



These are serious reasons for concern; they threaten society as we know it with the risks of mass loss of life, of culture and of states, of ecosystems and of biodiversity, of forced migration, and damage to health. It is harms such as these which make climate change such a pressing issue of justice. The 2°C target is based on the idea that these risks are unacceptably high once 2°C is surpassed.

These harms raise further questions of justice when we consider the fact that many of them are predicted to befall vulnerable communities who have contributed little to the problem of climate and lack a strong adaptive capacity (Gidley, *et al*: 2009). These features of climate change raise questions of justice with which philosophers have engaged. The climate justice literature is roughly as old as policy-makers' awareness of the problem of climate change, with Dale Jamieson (1990, 1991, 1992, 1996, 1998), Henry Shue (1992, 1993, 1994, 1995a, 1995b, 1996, 1999a, 1999b), and John Broome (1992), publishing on questions of climate justice since the early 1990's.

One of the prominent debates within the early literature on climate justice concerns which principle appropriately determines who should bear the burdens of responding to climate change: the polluter pays principle, the beneficiary pays principle, or the ability to pay principle (Shue, 1993, 1999b). Despite these principles offering distinct theoretical arguments regarding who should bear the burdens, they appear to converge on the same type of agents (developed countries), as the ones who should bear the burdens of acting on climate change. Sophisticated works have now informed the question of how to distribute the burdens of avoiding the harms from climate change. For example, Simon Caney (2012) has provided an egalitarian account of how these burdens should be distributed, while Mollendorf (2014) provides a specifically Rawlsian inspired account of how this should happen. For example, they both consider the question of how carbon emission should be distributed, in other words, who should be entitled to emit? One of the key lessons from this literature is that the harms of climate change should be understood as issues of justice, because they interact deeply with questions of distributive justice.

Distributive justice is not the only area of justice that interacts with climate change. For example, the idea of creating a procedurally just response to geoengineering will be central to the arguments in chapters 5 and 6 of this thesis. More generally, our responses to climate change raise pressing questions about how the relevant agents should be included or represented in decision making. This point has been made vividly by Marion Hourdequin (2012, 2016, 2018, and 2019) as well as being present in the work of Caney (2018) and Calles (2019a).

In 2004 Gardiner observed that moral philosophers had said very little about climate change. Since then, this has changed; the contribution of philosophers has resulted in the identification of a range of moral problems related to climate change. For example, in 2005, Walter Sinnot-Armstrong published a paper arguing that individuals do not have a moral obligation to act on climate change, by performing actions such as restricting their carbon emissions (Sinnot-Armstrong, 2010) . This has sparked a debate about whether individual emissions can be identified as harmful, and the nature of individuals' climate-related obligations (Broome, 2012, 2016; Cripps, 2013, 2016; Lawford-Smith, 2016).

The question of what we owe future generations in light of climate change is particularly tricky. There is the simple fact that harms from greenhouse gas emissions compound over time. So not only are we imposing a risk on future generations, but we are imposing a greater risk on

them than the one inflicted on ourselves. It is intuitive that the imposition of great and unwarranted risks constitutes an injustice. Yet the literature on future generations shows that this is more complicated than it may first appear. There is a famous debate about whether it is really possible to harm future people (Parfit, 1984; Parfit, 2010). This is followed by a somewhat perplexing debate on whether the welfare of future generations should be discounted. The perplexing nature of the discounting discussion is expressed by in some works of philosophers argue that position that discounting simply should not happen and that is grounded in a mistaken conception of utilitarian ethics (Moellendorf, 2004; Roser, 2009; Caney, 2014). This is countered by many technical arguments (Arrow, 1999; Broome, 1994; Dasgupta, 2012; Stern, 2007) one of which is the point that discounting welfare is not due to a presentism bias, but due to the need to account for uncertainties, such as the uncertainty of there being any humans in 500 years time whose welfare can be affected by our actions (Stern, 2007: 53). The point is that the idea of helping future generations is obscured by a fog of highly technical arguments.[4]

More recently the climate justice debate has concerned questions such as climate migration, criminalisation of emissions, loss and damage, reform of climate governance institutions and the role of ideal and non-ideal theory.

One of the debates about climate justice is that of the appropriate normative standard to use when considering climate change. This is one of debates with which is directly relevant ot this project. A cost-benefit style of analysis is adopted in prominent works such as The Stern Review which seek to evaluate responses to climate change. However, cost-benefit analysis can be ethically unsatisfactory (Pearce, 1998; Broome, 2012; Gardiner, 2016: 77). For example, it can invite trade-offs which do not make sense, by comparing all relevant considerations in monetary terms regardless of whether money is the appropriate metric to capture their value.[5] In response to the use of these cost-benefit frameworks, Simon Caney proposed that we should adopt a minimal rights standard for considering the problem of, and responses to, climate change. The minimal rights which Caney identifies are those of the right to life, health, and subsistence.[6] The idea is simple, which is that even this minimal set of human rights is capable

[4] Chapter 6 of this thesis does engage with part with the question of including future generations without the need to engage in these technical debates.

[5] The problems of a cost benefit analysis are expanded on in section 2.2 of this thesis.

[6] Caney also adopts a minimal understanding of these rights, For example the right to health is understood to mean agents have a human right that other people do not act so as to create a serious threat to their health, as oppose to much more expansive understandings of the right to health (Caney, 2010: 167).

of guiding our thinking about climate change and its responses. In those circumstances under which climate change violates these minimal rights, we can understand that something particularly problematic is happening.[7] Furthermore, the same is true of our response to climate change; they should not violate this minimal standard.

This minimal rights framework will be the normative standard which this thesis adopts. There are two reasons which, when combined, make the case in favour of adopting this standard: the benefit of a minimal standard, and a minimal normative standard. The benefit of a minimal standard is that the content of that standard is minimally controversial. In other words, people with a range of otherwise diverse moral commitments should be able to accept the content of the standard.[8] The analysis which uses this standard should not suffer from the concern that the content of the standard is odd or unreasonable. However, a minimum alone is not sufficient for the purposes of a thesis which is interested in ethical governance; for that we require the minimum to be an ethical one. This is where the importance of the standard being normative emerges. By having a normative standard, we can assess how just a situation is according to that normative criterion. The normative standard used here is that of rights. An appeal of a rights standard is that they are an expression of interests which are worthy of protection.[9] By having a minimal account of these rights, therefore, we can understand a situation as minimally just when this criterion is fulfilled. Moreover, when it is violated, we can understand a situation as failing to be even minimally just.

There are, of course, concerns about the minimal rights position. One of them is simply that it is too minimal, and it therefore misses important interests which are worth protecting. A response to this concern is that it does not make sense to think of protecting rights beyond this minimum if this minimum itself is not protected. For example, protecting one's right to education does not make much sense if rights to subsistence, health and life are violated. This response is grounded in Henry Shue's understanding of what a basic right is, which are those rights which need to be protected in order for other rights to be enjoyed (Shue, 1980). This provides another reason for focusing on this minimum, for it is necessary that these rights are protected in order for any more demanding standards to be fulfilled.

---

[7] From figure 1 we can see that the global mean temperature increases of 2 degrees or more will result in the violation of these minimal rights.

[8] The minimal standard in this thesis is uncontroversial, this can be seen by the lack of serious arguments against the idea of a right to life, health or subsistence. However, there are arguments about whether rights are the proper way to understand what is being protected here.

[9] I am using an interest based theory of rights (Shue, 1980), although there are theories of rights exists such as to protect autonomy or choices.

We find ourselves in a situation where there appears to be an unjust imposition of a grave injustice by a failure to prevent more than two degrees of warming. After more than twenty years of climate negotiations, there is little cause for optimism that the two degrees target will be achieved. This leads to a form of reasoning which asks what other options are available if our mitigation efforts are inadequate to prevent climate change. Some people, such as Paul Crutzen, have replied by considering geoengineering as a response to climate change (Crutzen, 2006).[10]

This proposal brings with it a great many questions. For example, is a geoengineered world preferable to one with extensive climate change? Are there other options available to us? Does geoengineering threaten the background conditions of life? Is geoengineering any more realistic than extensive mitigation? These are the types of questions which will be explored in my thesis. For the moment it is sufficient to say that geoengineering is being considered in response to the threat of climate change. This naturally raises the question: what is geoengineering?

[10] Crutzen's paper is noteworthy in part because it broke a perceived taboo on publishing about geoengineering.

1.1 What is geoengineering?

Before being able to ask questions about the ethics, governance or science of geoengineering, it is prudent to ask what geoengineering is. This thesis will focus on a specific form of geoengineering: stratospheric sulphur injection (SSI). But prior to introducing SSI, it is useful to have a more general understanding of geoengineering. A prominent definition of comes from a 2009 report by the UK's Royal Society on geoengineering, which defines it as 'the deliberate, large-scale manipulation of the planetary environment in order to counteract anthropogenic climate change' (Shepherd *et al*., 2009: 1). Geoengineering is clearly a broad term, for there are many different types of ways in which this 'manipulation' of the climate can be achieved.

The Royal Society report divides geoengineering technologies into two categories. These are carbon dioxide removal methods (CDR) and solar radiation management (SRM) (Shepherd *et al*., 2009: xi). As the name implies, carbon dioxide removal methods focus on the extraction of carbon from the atmosphere. This already happens naturally: think of a tree's relationship to the carbon cycle. Hence CDR either speeds up a naturally occurring process or it requires the implementation of new technology to remove the carbon. SRM focuses on either increasing the Earth's albedo or diverting solar radiation.[11] Diverting solar radiation is something that would seem much more at home in a science-fiction situation, given that discussions of SRM solutions often include putting sulphur aerosols into the atmosphere or placing mirrors in space. However, SRM can also involve much more down-to-earth actions such as painting buildings white to increase their ability to reflect the sun's rays (Shepherd *et al*., 2009: 23).

1.1.1 Terminological considerations

The Royal Society's understanding of geoengineering has been subject to much scrutiny. This is probably a quite natural and healthy practice. The report is one of the first of its kind on this topic, and it is on an issue about which there are strong reasons to be concerned from an ethical perspective.[12] Here I will raise a few of the most important concerns with the Royal Society's understanding of geoengineering. The concerns I wish to draw attention to are those of Dale

[11] Albedo refers to the reflective properties of a surface.
[12] These causes for ethical concern are explained in much more detail later in this chapter and throughout this thesis.

Jamieson (2013; 2014) and Stephen Gardiner (2011), who are both established writers on the ethics of climate change. I will then contrast this to the position of Clare Heyward who shows how the terms are helpful when applied correctly (Heyward, 2013).

Jamieson's concern is whether it makes sense for geoengineering to be a category of action which is distinct from mitigation or adaptation measures. For example, the Royal Society report includes afforestation as a type of CDR. However, the Inter-governmental Panel on Climate Change (IPCC) categorises afforestation as a type of mitigation (Jamieson, 2013: 528).[13] Jamieson makes this point to show that there is confusion about what we call geoengineering and what we call mitigation.

He pushes on this overlap and argues that we need to change the name of geoengineering and replace it with two separate categories of action related to either abatement or solar radiation diversion, abatement meaning any action which reduces GHG emission. [14] If these approaches are made distinct categories of action alongside mitigation and adaptation, Jamieson thinks that all climate change responses could fit neatly into one or other of these categories (Jamieson, 2014: 207). For example, according to Jamieson, afforestation would then be a clear case of an abatement strategy. Therefore Jamieson's concern is not only with the Royal Society; his argument about afforestation being a form of abatement means that he also disagrees with the IPCC's understanding of mitigation. Despite Jamieson's concern about the name 'geoengineering', the term still dominates the literature on this topic. For example geoengineering is mentioned in the most recent report by the IPCC and it is the subject of many governance reports as well (IPCC, 2014: 89; Morrow, 2017).[15]

Moving away from Jamieson, Stephen Gardiner's concern with the Royal Society report is that he doesn't believe there is any need to have a sharp distinction between CDR and SRM (Gardiner, 2011: 345). This is because we can imagine some very strong forms of CDR which would raise many of the same issues as SRM. For Gardiner, the type of questions we ask about different geoengineering technologies depends on the effects of those technologies. There is no reason why CDR and SRM technologies should necessarily have a different set of effects.

[13] The IPCC is part of the United Nations Framework Convention on Climate Change (UNFCCC). The IPCC is tasked with providing policy-makers with the most up-to-date accounts of climate science and the impacts of climate change.

[14] For Jamieson abatement is distinct from mitigation because mitigation focuses on efforts to initial release of the greenhouse gases emissions in the atmosphere, whist abatement focuses on reducing the quantity of greenhouse gases which are already in the atmosphere (Jamieson, 2014: 207).

[15] I am aware there is something unsatisfactory about this logic, in that using a word because it is dominant only perpetuates its dominance – particularly if enough people follow the same logic.

Therefore this distinction has no value other than giving a different name to a different set of methods which produce the same effects.

In contrast to both Jamieson and Gardiner, Heyward provides an account of how the categories of CDR and SRM should be used (Heyward, 2013). Heyward's argument is that CDR and SRM both exist at points on the continuum of response to climate change. Importantly they should be kept distinct because they have different aims, the aim of CDR is that of '[a]voiding a given level of atmospheric GHG [greenhouse gas] concentration', and the aims of SRM are '[a]voiding global average temperature increases' (Heyward, 2013: 25). Heyward's point is that the names are fine, in fact useful, as long as we are sensitive to the distinct aims of CDR and SRM.

Gardiner's and Jamieson's concerns seem to be similar, insofar as they are both concerned with the fact that the terminology around geoengineering does not really capture what is important. For Jamieson, the issue is with the word geoengineering itself, and for Gardiner it is about dividing geoengineering into two separate categories. It should be noted as well that this is not the extent of either of their concerns with geoengineering; in fact, this may be considered among their least forceful objections. Nonetheless I think this divide and the source of this division can be identified. Jamieson and Gardiner seem to have different criteria in mind when considering the name geoengineering. Jamieson seems to be seeking a term which is descriptively accurate of how geoengineering technologies interact with other systems. Whilst Gardiner seems to be seeking a name which does not obscure the ethical considerations at stake. On the other hand Heyward appears to think that descriptive accuracy can be achieved as long as we are clear about the aims of CDR in contrast to SRM technologies. It is interesting to note that these different criteria have resulted in vastly different conclusions given that they are clearly related. Descriptive accuracy is important for highlighting normatively salient features of a technology. I mention these concerns to show that the term geoengineering is one which we should be cautious of.[16] Despite this, there are strong reasons to stick with the term, not least because it dominates the literature on the subject. The term dominates the literature on the subject, for example, geoengineering is mentioned in the most recent IPCC report and is the subject of many governance reports (IPCC, 2014: 89; Morrow, 2017).[17] Therefore the term will

[16] Such caution can be seen by the Carnegie Climate Governance Initiative, which removed the term geoengineering from its name due to the confusion that it caused. It was previously known as the Carnegie Climate Geoengineering Governance Initiative (Pasztor, 2019).

[17] I am aware of something unpleasant about this logic, using a word because it is dominate, would only perpetuate its dominance. Particularly if enough people use the same logic.

be used in this thesis though hopefully with the caution that it deserves. This is not a significant challenge in part due to the vast majority of the thesis focusing on a specific type of geoengineering, which is explained towards the end of this chapter.

1.1.2 Geoengineering: a new idea?

Geoengineering may sound like an idea which belongs to a futuristic dystopia. However, the idea of trying to control the weather has been around in some form for millennia. That is, the idea of trying to control the weather. We can trace the notion of governments having an interest in weather control to at least the beginning of the nineteenth century. For example, the US government offered funding for experiments which explored the idea of weather control for most of the 19th century (Fleming, 2010: chapter 3).[18] An interest in controlling the weather can also be seen throughout the history of the Soviet Union. Lenin provided a philosophical foundation for this by reasoning that humans were on the precipice of being able to master nature (Fleming, 2010: 198). Subsequently, in 1948 Stalin announced a great plan for – among other things – weather and climate control.[19] We can therefore see that the idea of interfering with the weather is not necessarily as novel as it is sometimes portrayed. Even the term Geoengineering is not as young as one may expect, the first use of the term 'geoengineering' can be traced to an academic publication in 1977 (Marchetti, 1977). That's thirteen years before we can claim that there was an international consensus on the threat of climate change, and even more years before we could claim a consensus on our failure to act on it.

[18] The idea of controlling the weather is much older even than this. For a historical overview, see James Fleming's *Fixing the Sky* (2010), in which he points to the idea of weather control going back to at least Greek mythology, when Phaeton is given control of the sun (Fleming, 2010: 16). There is also a lesson of hubris from this story – hubris being an idea which is explored in chapter 2 of this thesis.

[19] This plan is consistent with the philosophy of Lenin from a few decades earlier (Fleming, 2010: 198).

1.2 Why there should be a formal governance regime for geoengineering: Real-world and ethical causes for concern

In this section I will explain why there should be a formal governance regime for geoengineering. The case for governing geoengineering which I will present is predicated on two conditions being met.[20] Firstly, that geoengineering is being taken sufficiently seriously in the 'real world', and secondly, that there are ethical concerns with geoengineering. Once these two conditions are met, there is a compelling case for geoengineering to be properly governed.[21] This section will proceed by establishing that both of these conditions have been met in the case of geoengineering

Prior to making the case for geoengineering governance it is worth heeding the warnings for the case for such governance should be made. These warnings are often about how adopting certain frames to justify geoengineering governance are problematic. Here I shall draw attention to two such concerns, those of Heyward (2015,) and of Fragnière and Gardiner (2016).[22] Heyward argues that we should not adopt a lens of exceptionalism when considering geoengineering as such lenses are simplistic and damaging to public discourse. The damage to public discourse is a product of the exceptionalist lens leading to agents overstating the virtues or vices of geoengineering (Heyward, 2015). Fragnière and Gardiner also providing a warning about the framing of geoengineering, one which can be understood as an argument that we need to avoid certain frames when discussing geoengineering and this includes making the case for governance. They argue that the plan B or insurance framing of geoengineering is misleading, for make geoengineering could more attractive than actually is by hiding important ethical considerations which accompany geoengineering. There are of course other framing concerns about SSI, and some of them are considered in more detail in chapter 2, yet these framing concerns are serve as a useful warning about how arguments in favour of SSI governance should not be made.

[20] This is not a general argument about the circumstances according to which something ought to be governed, although these conditions may be relevant in many other cases as well.
[21] Janos Pasztor (2017) points to similar conditions when calling for geoengineering to be governed.
[22] Chapter 2 will draw attention to the other framing concerns which exist about geoengineering.

1.2.1 Geoengineering is being taken sufficiently seriously in the 'real world'

By stating that geoengineering is being taken sufficiently seriously in the 'real world', I mean that geoengineering is genuinely considered to be a plausible response to climate change among certain relevant agents. In this sub-section I shall point to the popularity of geoengineering among a diverse group of actors, as well as the different pushes for geoengineering to be properly governed and researched further. These types of activities are indicative of geoengineering being taken seriously by at least some agents.

Geoengineering is popular among a range of agents. It is popular among certain economists, certain scientists and certain politicians (Barrett, 2008; Gingrich, 2008; Levitt and Dubner: 2009: 194-202; Keith, 2013). To elaborate, there are various groups of scientists – most prominently a group at Harvard University led by David Keith – who take the idea of geoengineering seriously as a potential response to climate change and are undertaking research into geoengineering technologies. (Below I provide more detail about a recent research project they have carried out.) The popularity of geoengineering among certain economists is illustrated by the title of an article by Scott Barrett called 'The remarkable Economics of Geoengineering', which makes the argument that certain geoengineering technologies are significantly cheaper than efforts at mitigation (Barrett, 2008).

This has resulted in some economists misrepresenting the virtues of SSI. For example, in a popular book called *Superfreakonomics*, climate change is portrayed as not being a problem at all due to the cheap solution provided by SSI (Levitt and Dubner, 2009: 194–202). Moreover, comparable errors have been made by politicians. In a blog post in 2008 Newt Gingrich (Gingrich, 2008) the former speaker of the United States House of Representatives argued against funding mitigation efforts because there is the cheaper alternative of funding geoengineering 'innovation'. In short, geoengineering is a popular idea among some sub-sets of certain groups who have, or try to have, influence on policy.

To elaborate on the claim that some scientists are taking geoengineering research seriously: in 2018 plans for a project called SCoPEx (the Stratospheric Controlled Perturbation Experiment) were published. Led by Keith, this project aims to perform experiments to provide additional information about the viability of geoengineering, via field tests (Tollefson, 2018). The intention is for this project to happen in 2020. The SCoPEx project is being carried out in the absence of any regulatory framework specifically designed to govern SSI research. Whilst the

project may not raise too many significant ethical questions in isolation, it is indicative of a willingness, and capacity, among certain actors to pursue research into and maybe the development of SSI technologies, independent of a governance structure that has been specifically designed to take account of geoengineering. Another such experiment is the SPICE (Stratospheric Particle Injection for Climate Engineering) project at Cambridge University, which proposed to pump 150 litres of water into the atmosphere, to observe how the pipe and water would be effected by various wind conditions (Cressey, 2012).[23]

SCoPEx and SPICE are illustrative of geoengineering research projects which have happened. Between 2008 and 2018 there was an increase in funding of geoengineering related projects from $1 million to $8 million, with most of these projects being based in the UK, US and Germany (Necheles *et al*: 2018). An interesting detail is that from 2008-2018 the vast majority of projects in Europe were government funded, whilst in North America the vast majority were privately funded (Necheles *et al*: 2018).[24] This shows us that geoengineering research is happening and of interest to actors in the public and private sectors across North America and Europe. Moreover there are small research programs in other continents such as Asia and Australasia.[25]

Given the popularity of geoengineering among certain relevant agents, it is hardly surprising that geoengineering has caught the attention of the wider political community. This can be seen in part by the numerous reports which are being produced on the governance of geoengineering. Between 2009 and 2014 at least ten governance reports were produced.[26] These reports can be produced at the behest of political actors. For example, the Royal Society report was produced due to a request for information from the UK House of Commons Science and Technology Committee. Recently, the real-world interest in geoengineering has moved beyond a desire just for reports. Switzerland proposed a resolution at the Fourth United Nations Environment Assembly calling for an assessment of potential governance frameworks and methods. Despite

[23] Although the absence of governance may not be the main reason why the SPICE project was cancelled, there were concerns about a conflict of interest over a patent held by a member of the team (Cressey 2012)

[24] It is challenging to be precise over exactly how much is spent, in part due to difficult questions as to what counts as geoengineering research in contrast to an atmospheric physics experiment.

[25] There were no noteworthy geoengineering programs in South America or Africa between 2008-2018, but this should not be mistaken for a lack of interest in the technology.

[26] Namely: The NOVIM Report, The Royal Society Report, The Congressional Research Service Report, The Bipartisan Policy Center Report, The Kiel Earth Institute Report, The SRMGI Report, The Wilson Center Report, The Ecologic Institute Report, The National Academy of Sciences Report, and The EUTRACE Report.

this resolution being rejected, it does indicate that a conversation is already taking place about how geoengineering can work.

The examples above are meant to illustrate that geoengineering already appears to be popular among certain actors and that moves are being made for geoengineering technologies to be researched and developed and also governed.[27] Yet this is not enough on its own to justify the need for there to be principles of geoengineering governance. For that to be the case, it also needs to be clear that there are ethical causes for concern with geoengineering technology.

1.2.2 Reasons for ethical concern

Specific reasons for ethical concern will be explored in greater detail in chapter 2 of this thesis.[28] For now it is sufficient to outline why there may be some ethical considerations with regard to geoengineering. As we have established, the climate is a background condition of life and geoengineering alters the climate, in ways that could have serious repercussions for basic rights. Therefore geoengineering alters a background condition of life. This leads to ethical considerations such as: should humankind have the power to alter the background conditions of life? If the answer to this is 'yes', then the next questions are: should this power be distributed evenly? Could this power be legitimate? Would this power change how humans understand the other background conditions of life? I shall therefore explore how the whole issue of geoengineering raises questions about significant ethical considerations such as responsibility and harm. And by doing this, I will show how geoengineering is indeed an issue of ethical concern.

Responsibility is an important notion in ethics. If we break something, we tend to think that we have an ethical duty either to fix it or replace it.[29] Hence there are questions about whether geoengineering undermines our common accounts of responsibility. If we accept that geoengineering is a less desirable option than climate change mitigation, then if the rich nations which refuse to mitigate were to offer geoengineering solutions instead, it would appear that

[27] This is not a complete account of all the reasons for real-world concern with geoengineering since one could also point to the poor framings of geoengineering which are prevalent in public discourse some of these framings are considered in chapter 2.
[28] The concerns in chapter 2 will be specific to the type of geoengineering this thesis focuses on:SSI. The purpose of this subsection is to outline a more general account of why we should be interested in geoengineering ethics.
[29] Responsibility is a more complicated idea than this, but this captures our intuitive concern about responsibility.

they were not fulfilling their responsibilities because they are failing to act in a way that would minimise the impact of the problem they have created. They are not adequately fixing or replacing what they have broken. Therefore, insofar as we value responsibility, we need to treat arguments in favour of geoengineering with caution.

Additionally, there is the appeal to one of the most basic of ethical considerations, which is the idea of doing no harm. Geoengineering is likely to be a harmful act, but we do not know the extent of the harm it will cause. It might lead to an increase in acid rain, or interfere with weather patterns, and it may well defer harm onto future generations, in ways that would violate their basic rights (Shepherd *et al*., 2009: 32).[30] This means that geoengineering technology is likely to create new harms which we do not currently experience as a result of climate change. Moreover, it is not going to prevent all the harms caused by climate change. For example, geoengineering technologies which divert solar radiation do nothing to prevent ocean acidification (Shepherd *et al*., 2009: xi).[31] The idea of geoengineering creating new harms and failing to alleviate old ones means that we have reason to find it ethically troubling. And this is only the beginning of the harm-based questions. For example, we might be worried about the distribution of harms from geoengineering or want to ask if they correlate to responsibility. These types of question are more complicated than I am suggesting here and will recur throughout the thesis. For the purpose of establishing whether geoengineering gives cause for ethical concern, it is enough to say that it creates some new sources of harm.[32]

1.2.3 A case for governance in light of these real-world and ethical concerns

Standing alone, none of these reasons may be compelling enough to take the governance of geoengineering seriously, but when considered together, the case is stronger than the sum of its parts. Considering the real-world reasons for the governance of geoengineering on their own, then just because something is popular, that is not necessarily a reason to govern it.

[30] The reason why geoengineering may defer harm onto future generations is complicated, but it is explained in the section on the 'termination problem'.

[31] The reason for this is that ocean acidification is a product of the carbon levels in the atmosphere. And SRM technologies (such as SSI) do nothing to decrease carbon levels. If anything, SSI may well incentivise high carbon levels. For an explanation of why there may be higher carbon levels, see my discussion of the moral hazard.

[32] There is another way we could make the case for geoengineering being an issue of ethical concern, which is simply to observe that there is a huge literature on the subject of climate change as an issue of justice. Consequently we would expect any prospective response to climate change to address these issues. It would be unusual if a response to an issue of justice did not raise any ethical concerns

Harmless (or minimally harmful) activities can be popular without the need for governance: the game of squash, for example. It is only when we consider the reasons to be ethically concerned about geoengineering as well that a case for governance becomes clear.

The same point is true with regard to ethics. Something being deemed an ethical consideration is not sufficient grounds for governing that thing. There are at least two reasons why this may be the case. Firstly, the issue could be deeply personal and not the type of question which is subject to governance. The second reason is that governance needs plausibility. For example, we could say that there is something deeply unethical about the destruction of other galaxies, yet this does not create an onus on us to legislate to govern or prevent this because there is currently no plausible way for us to destroy other galaxies.

If we only focus on issues which have both real-world and ethical reasons to take into consideration, then we are much more likely to have a case for governance. This may not always be so, of course, and I am sure it would be possible to imagine a case for governance without any real-world or ethical causes for concern. Nonetheless, with regard to geoengineering, it seems that the case for governance is more compelling when we consider the ethical and real-world causes for concern in combination. Taken together, the fact that geoengineering is popular among certain actors provides a real-world cause for concern, whilst the potential for inflicting additional harm indicates an ethical cause for concern. We then have a popular idea which may be harmful, and this clearly creates a need for good governance of the technology. [33] Moreover this case for governance does not fall foul of warning of exceptionalism or plan B framings of the need for geoengineering governance.

[33] There is an alternative type of argument which can be offered in favour of geoengineering governance, and that is the anticipatory governance argument. The anticipatory governance argument emphasises the importance of foresight, engagement, and integration (Foley, *et al*, 2018: 228). This approach to governance is proposed for governing other types of emerging technologies, such as synthetic biology and nanotechnologies (Gorman, 2012). This approach also has some appeal due to it apparently problems which are associated with precautionary approaches to governance (Foley, *et al*, 2018: 228. This thesis is not deeply grounded in the anticipatory approach to geoengineering governance, in part because there is not much literature on how this approach would look in the context of SSI governance, with the exception of Foley, *et al*, (2018). I understand my approach as one which may be complementary to anticipatory approach, but not bound by it.

Understanding geoengineering governance: what is it and what has been proposed?

In this section I shall offer clarity about geoengineering governance and the level at which it can occur. The clarity which I will provide is about the levels at which geoengineering governance can occur. Governance is a broad term, it captures the ways in which rules or guidelines are used to influence the research, development, or deployment of geoengineering (Chhetri *et al*, 2018: 3; Reynolds, 2019: 6). There are a range of ways in which such governance can be achieved. These range from the establishment of norms which regulate behaviour, to the establishment of a centralised authority, which has the power to use coercive force to make agents obey a set of rules and regulations. This range can be clearly seen in Reynolds 2019 review of geoengineering governance proposals, in which he identifies at least 5 categories geoengineering governance (Reynolds, 2019: 15-17).

The first level of governance which Reynolds identifies is that of the deliberative intergovernmental institution where such an institution has the power to make important decision about the direction of geoengineering research, development, and deployment. (Reynolds, 2019: 15-16). For example, an institution which is comparable to the UNFCCC[34], the idea of governance at this level is argued for by Albert Lin (2009). Looking beyond the UNFCCC Adam Abelkop and Jonathan Carlson (2013) as well as Ralph Bodle (2014) have argued for other international institutions to govern geoengineering. The distinctness of this level of governance can be seen when comparing it to the second level of governance which Reynolds identifies. The second level 'of governance also concerns intergovernmental institutions. However, in this case their role is only that of facilitating certain agreed upon actions between states, such as sharing information and resolving disputes, not that of coordinating all or any geoengineering research, development or deployment programs (Bodansky, 2013; Armeni and Redgwell, 2015; Reynolds, 2019:16).

The third level of governance which Reynold identifies is that of a small groups of states together to govern SSI (Reynolds, 2019:17). There is variation between scholars about who these small group of states are and what powers they would have (Benedick, 2011; Lloyd and Oppenheimer, 2014: Reynolds and Wagner: 2019), yet the general theme appears to be that the number of states required is between 25-30 and they would have advisory power about SSI

[34] Which is an environmental treaty from 1992, the aim of this treaty is to coordinate global responses to climate change.

decisions. Yet this subject to the concern that if the powers are of only advisory would this be really be a site of governance which is fit for purpose of adequately addressing the ethical challenges which geoengineering creates. If the answer is no, then such governance appears at best inadequate and at worst a distraction and waste of time when time is limited. Yet if such governance doe have the power to make meaningful decision about SSI then it runs into a legitimacy problem due to the decision being made by a small minority of states.

The fourth level of governance is not focused on states at an international level (Reynolds, 2019:17). Instead the focus is on developing appropriate norms and codes of conduct. This has been proposed by Edward Parson (2017). Parson's argues that this focus would be helpful for the further geoengineering governance. Whilst there is an appeal to this approach, namely that it does not seem to suffer from the challenges of having to navigate or create an intentional governance institution, its not without it its problems. An important concern with the norms based approach is simply whether or not norms are really reliable enough to act as be a source of governance of geoengineering in which we should place our trust.

The FCEA report offers a broader understanding of national level governance, pointing to how nations already have regulatory frameworks which could capture much geoengineering activity at the national level (Chhetri *et al,* 2018: 23). For example, environmental law may restrict the possibility of field tests. Moreover if such law is inadequate national governments can develop their own laws to regulate geoengineering activates within the boundaries of their state. Finally you have those who hold that geoengineering should not be used, therefore governance should occur via a moratorium (Reynolds, 2019: 17). This view is held for variety of reason such as fears that it cannot be compatible with democracy (Szerszynski 2013;Hulme, 2014),[35] that it raises security problems (Cairns and Nightingale, 2014), and it violates international law (Winter, 2011).[36]

This shows that when thinking of governance one could vary from thinking of high level international institutions, to norms at the regional level or a moratorium on geoengineering research, development, and deployment. Therefore geoengineering governance of some form is already being practiced. There are norms which inform scientific practice when

[35] Counter arguments to the view that geoengineering is incompatible with democracy can be in found in Heyward and Rayner (2016) and (Horton (*et al*; 2018).

[36] One might think that if geoengineering is problematic then the moratorium position allows you to avoid these problems, this is far from true, this is demonstrated in chapter when the risk-risk trade off frame is explained.

geoengineering is being researched, and there are domestic laws which act as a constraint on the remit of research projects (Reynolds, 2019: 7).

It should now be clear why we have an interest in the governance of geoengineering and how geoengineering governance can be understood. This allows us to explore how the principles which have been proposed would govern geoengineering. To do this I will proceed by introducing perhaps the most famous proposed principles of governance, the Oxford Principles. This will be followed by considering some of the concerns which have been raised about these principles. Finally, this section will end by considering the work of Morrow (2017) who provides a summary of the principles which have been proposed from 10 repots, importantly drawing attention to where there is a consensus on these principles. The subsection should provide a picture of what principles exist to govern geoengineering, and make clear where in this picture the thesis will end up sitting.

The Oxford principles are among the most influential set of principles proposed for the governance of geoengineering (Rayner *et al*, 2013: 500). They are the product of an inquiry by the UK House of Commons Select Committee on Science and Technology, launched after the Royal Society report on geoengineering, in order to find some principles to govern geoengineering. These principles have a broad scope because they were required to cover the governance of all geoengineering technologies.

Perhaps of the first thing to note about these principles is that they were designed to cover the governance of geoengineering from research and development to deployment (Rayner *et al*, 2013: 500). This means that they are designed to cope with a wide variety of problems. The principles are as follows:

‘1) Geoengineering to be regulated as a public good

2) Public participation in geoengineering decision making

3) Disclosure of geoengineering research and open publication of results

4) Independent assessment of impacts

5) Governance before deployment’ (Rayner *et al*, 2013: 502-503).

The function of these principles is to act in the way legal principles tend to act (Rayner, 2013: 504). This means they are not designed for particular cases, but tot be interpreted in the light of the specific cases. When discussing geoengineering, we can replace the word case, with geoengineering technologies.

The Oxford principles are inform norms and procedures which are used to determine how geoengineering activates should take place, For example the SPICE project did have public engagement in order to conform to the second Oxford principle (Rayner *et al*, 2013: 509; Pidgeon *et al*, 2013, 451).

However, whilst the Oxford Principles are broadly endorsed by many agents, such as UK government and by members of the international scientific community at the Asilomar Conference (Rayner *et al*, 2013: 500), they are not without their critics. A recurring concern amongst these detractors is that the principles operate at high a level of abstraction (Nature 2012). Although this point seem to a product of people failing to appreciate the purpose these principles (Rayner *et al*, 2013: 504; Gardiner and Fragnière, *2018: 145*). They are designed to operate at a high level of abstraction, and there are benefits of their doing so, such as their not requiring to rule on specific issues of geoengineering governance. This is particularly important as their do so may mean ruling on issues prior to the existence of the information necessary to make such rulings. Gardiner and Fragnière (2018) do provide a critique of the Oxford Principles. They are concerned about the frames produced from these principles and offer an alternative set of principles named the 'Tollgate principles'. These are meant to be address the concerns they raise about the Oxford principles. In short their fear is that the Oxford principles leave open the possibility of ethics being neglected. For example, they argue that the 1st Oxford principle, which frames geoengineering as a public good, obscures the fact that there are important interests at stake for both the current generation, future generations and non-human nature (Gardiner and Fragnière, 2018: 152). Their argument is that the first Oxford principle presents geoengineering as a good for all, and consequently obfuscates the fact that there are different agents' interests at stake which we should consider when certain decisions about geoengineering are being made. Consequently, the ethical considerations are harder to spot. This is true if one just reads the principles, although if one reads the Oxford Pinciples paper in full the legitimacy of this concern may be diminished, for the paper is explicit that the first principle 'requires consideration of global and intergenerational justice' *(*Rayner *et al*, 2013: 505).

It is not my intention to settle the debate between the Oxford and the Tollgate principles, only to give the reader a sense of what principles of governance for geoengineering have been proposed. Whist the Oxford and Tollgate principles are among the most noteworthy principles proposed for the governance of geoengineering they do not stand alone. David Morrow has produced a comprehensive survey of the views expressed in ten governance reports on geoengineering. Morrow identifies a number of principles of geoengineering governance from these reports and, depending on the degree of support these principles have across the reports, categorises them according to whether there is 1) universal consensus, 2) broad agreement or 3) unresolved questions about the particular principle. Table 1 is based on the information contained in Morrow's survey and is designed to demonstrate in a simple format which categories Morrow places particular principles into. Each of the principles listed is quoted from Morrow (2017, 6–12), who uses climate engineering (CE) as an alternative name for geoengineering. I shall not consider these principles in a point-by-point manner, since many of the points listed do not play a role in this thesis. Instead, I will focus on some of the points outlined in Table 1 as well as the ten reports on geoengineering governance, as appropriate.

Table 1. Geoengineering Governance Principles as Categorised by Morrow (2017, 6–12)

Universal consensus	**Broad agreement**	**Unresolved questions**
'If CE is ever deployed, it should be in addition to traditional mitigation and adaptation measures'	'Governance should be proactive rather than reactive'	'What are the objectives of CE governance?'
'Existing laws and institutions provide partial governance of CE but additional governance mechanisms are needed'	'Governance arrangements should be flexible and adaptive'	'Which institutions should take on which functions of CE governance? How should these institutions and the various governance mechanisms at their disposal relate to one another?'
'A general moratorium on CE research is inadvisable at this time'	'For now, informal, soft-law approaches to CE governance are better than formal, hard-law approaches'	'Should there be an "allowed zone" for research? If so, how is it to be defined?'
'Governance structures should encourage international cooperation and coordination on CE research'	'Governance must strike the right balance between legitimacy and effectiveness'	'How should transparency be operationalised?'
'If research does proceed, transparency and openness are critical'		'What form should public engagement take and how should it shape CE research and governance?'
'Public engagement is desirable'		'What role should precaution play in CE governance?'

Despite all the reports on geoengineering governance and all of the above principles, it would be a mistake for those who care about ethics to assume that we have an appropriate ethical apparatus for the governance of SSI. At the very least, it would be worth paying attention to some important ethical concerns about geoengineering and asking what a governance arrangement would look like which addressed this concern adequately. And asking too whether the geoengineering governance principles in Table 1 are adequate to address our ethical concerns.

The aim of this introduction has been to present the case for why we might be interested in SSI being governed ethically and to introduce the reader to the idea that there are many proposed principles for the ethical governance of SSI and geoengineering more generally. The question now is: where does this leave us? There is a plethora of questions one could ask at this point. One could still question whether it is even appropriate to be having a discussion about the principles of SSI governance. Or one could choose some of the proposed principles and try to

evaluate critically whether they are actually desirable. Or indeed, one could take feasibility constraints very seriously and decide that these principles are implausibly idealistic.

1.3 What the thesis will do

My thesis adopts a different approach, however, because it is premised on the idea that a just regime for geoengineering can be designed to address certain ethical concerns about geoengineering. The thesis will focus on three specific concerns – the moral corruption concern, the moral hazard problem, and the question of how to include future generations – and it will provide some ways to think about whether the proposals for SSI governance might be capable of addressing these important ethical concerns.

To achieve this, the thesis will proceed in the following way. Chapter 2 will provide the reader with the information they need to understand the thesis, to do this the chapter will introduce the reader to other ethical concerns with geoengineering, in order for the reader to be able to situate the concerns which this thesis focuses amongst the broader sets of concerns. Secondly the chapter will introduce two important frames which are used throughout the thesis, those of the risk-risk trade off and the prefect moral storm frame.

Chapter 3 focuses on the problem of moral corruption. In short, the moral corruption fear is that agents will deceive themselves into thinking that SSI is ethically required or permissible in situations where it is not. This has a range of problematic implications, ranging from poor quality public dialogue on SSI to the potential for ethically problematic climate policy. The approach of Chapter 3 is to consider how transparency and accountability could serve as tools to address moral corruption. This investigation leads on to a criticism of how the principles of transparency and accountability have been articulated in geoengineering governance reports. Specifically, that these reports tend not to appreciate the necessary relationship between the two concepts if accountability is to function well and genuinely address problems such as moral corruption.

Chapters 4 and 5 focus on the moral hazard problem. In short, the moral hazard concern is that there will be a reduction in mitigation efforts due to SSI, and that this is bad, either because it increases the intensity of the climate change which will occur or because it increases the likelihood of SSI being used as a means to prevent such intense climate change. Chapter 4 focuses on how the moral hazard complaint should be understood, drawing on both how the complaint is articulated in geoengineering governance reports and the philosophical literature. The chapter shows that the moral hazard complaint has yet to be formulated in a way which is fit for the purpose of being addressed via principles of governance. It rectifies this by providing

sufficient clarity about what the moral hazard complaint is, and this provides a foundation for being able to judge whether the complaint is addressed by a specific set of governance principles. Chapter 5 takes up part of this challenge by considering an intuitive solution to the moral hazard problem, namely, secrecy. This chapter has the descriptive aim of showing how secrecy could address the moral hazard problem, and this is then followed by a normative component of rejecting secrecy as a desirable principle for addressing the problem.

Finally, Chapter 6 considers a different type of problem, which is how future generations could be represented in an SSI governance regime. The reason for considering this question is that if we accept the need for accountability in this context, then it appears that a just conception of accountability would require an SSI governance institution to be accountable to future generations in some form as well.

There is variation in the approaches adopted in these different chapters. I understand this variation as a product of these chapters being interested in different kinds of questions. Chapters 3 and 5 are concerned with similar question of trying to understand if particular principles such as accountability and secrecy are capable of addressing particular concerns with geoengineering. Both of these chapter start by considering these principles at a theoretical level, but to judge their appropriateness for engaging with problems, both chapters consider what these principles could look like in practice. This is done by looking at examples of accountability and secrecy mechanisms in operation.

This is in contrast to the fourth chapter which is conceptual in nature, it is focused on understanding how the moral hazard can be understood, therefore it does not have this same kind of real-world practice to observe and draw upon. Instead the chapter focuses on the theoretical literature about the hazard, and the hazard is being used in reports on geoengineering governance to present an understanding of the hazard. Finally chapter 6 is interested in a different type of problem. I understand chapter 6 as asking an applied type of question, which goes if we accept a principle which requires the representation of future people how could that principle appropriate in geoengineering governance. This is a distinctive question from the type which has been asked in other chapters, and whilst the chapter can draw on certain mechanisms which are in the real world like that of second chambers it is considering a mechanism of which there is a distinct lack of in the real world, so cannot there are case of this mechanism to draw upon in the way that chapters 3 and 5 draw on the mechanism of accountability and secrecy. Therefore chapters 3 and 5 can be understood as both asking questions about the

appropriateness of a principle, chapter 4 as understanding the nature of a problem, and chapter 6 as being interested in how a principle can be applied.

By focusing on these issues, the thesis will provide some tools for judging whether a specific set of geoengineering governance principles established by a particular institution is well placed to address the moral hazard and moral corruption concerns. This is indicative of how just this institution might be, given that these are both strong causes of ethical concern in relation to SSI. Yet these questions do not exist solely in the domain of geoengineering governance reports. There is also a political theory literature which seeks to engage with questions such as these, and indeed, these concerns were first articulated in theory. Consequently, this thesis also contributes to the political theory literature on governing SSI. The value of the thesis therefore is that it contributes to at least two valuable endeavours: firstly, judging whether principles of geoengineering governance do actually address certain ethical concerns, and secondly, contributing to the development, understanding and solutions for ethical concerns as expressed in the political theory literature.

1.3.1 Narrowing the scope of the thesis: clarifying the type of geoengineering

As has been explained, geoengineering can be performed in many different ways. The options vary from reforestation to putting mirrors in space. Given these two extremes, it may already be clear that these different methods raise quite different questions of governance. Consequently, given that different technologies create different questions of governance, it is necessary for me to choose a specific technology to focus on if I am going to establish some specific principles of governance. In order to make this choice, I propose to draw on the criteria set out in the previous section and show that one particular type of geoengineering technology excels at meeting the condition of both real-world and ethical concerns with geoengineering. That technology is called solar stratospheric aerosol injection (SSI).

SSI is the idea that geoengineering is possible by putting sulphur particles in the stratosphere. These particles have reflective properties, they scatter the incoming solar radiation, and consequently reduce the total amount of solar radiation which passes through the stratosphere. This will have a cooling effect, because solar radiation is a source of heat. Our knowledge of this is partly based on how we understand the effects of volcanic eruptions (Shepherd *et al*.,

2009: 29). When volcanoes erupt, they send sulphur into the stratosphere and this has a cooling effect.

Whilst SSI is thought about as a way to reduce the global temperature it does have other impacts as well .Modelling has shown that SSI could have the following effects: Drought in Africa or Asia, Ozone depletion, continued ocean acidification, impacts on tropospheric chemistry, and raid warming if stopped (Robock, 2016).[37] The extent of these impacts is dependent in part on the quantity of sulphur which is injected into the stratosphere (Irvine *et al*, 2016: 95).

SSI has some interesting features. It is relatively cheap. It could cost as little as $1 billion per year to deploy (Keith and MacMartin, 2015). The relative cheapness is clear when the cost of SSI is compared to the cost of mitigation efforts which could be $200-350 billion per year by 2030 (Ritchie, 2017).[38] Callies helpfully shows the relative cheapness of SSI by reminding us of how much damage can be done by unmitigated climate change, which Nicholas Stern projects would cost 5% of global GDP and William Nordhaus predicts it would be 2.5% of GDP (Callies, 2019a: 19).[39]

An additional interesting feature of SSI is that it starts having an impact on the global temperature within weeks of being deployed, this is fast by many standards, and importantly it is relatively really fast. This point is clearly made in the Royal Society report on geoengineering, which identifies SSI as the being amongst the fastest forms of geoengineering to have an effect on the global temperature (Shepard, 2009: 48).[40] The cost and efficacy of SSI are appropriately described by Daniel Callies as SSI's 'great merits' (Callies, 2019a: 7).

I will now show how SSI fits the condition of real-world concerns. SSI has always had a prominent place in discussions about geoengineering among policy-makers and scientists. It had the privilege of being the example of geoengineering used in a publication by Paul Crutzen, which is perceived as breaking the taboo on the subject (Crutzen, 2006), and it has kept its place as one of most discussed forms of geoengineering. This can be seen in the work of people such as David Keith, who has written a book which makes the case in favour of SSI, and Mike Hulme, who has written a book arguing against SSI (Keith, 2013; Hulme, 2014). The economist

[37] Some of these effects are considered in more detail in chapter 2 such as the raid warming which would occur if SSI were stopped.
[38] Although this understand of the cost of mitigation is one which does not account for the co-benefits of mitigation
[39] It is important to note that even the conservative number of 2.5% of GDP is significantly high than the cost of mitigation which are not even 1% of global GDP (Ritchie, 2017)
[40] It can also have an impact much on global temperature much faster than mitigation efforts, due to life time of carbon in the atmosphere, this was explained in the first part of this chapter,

who engage with geoengineering such as Barret and Levitt use the case of SSI to make their arguments, for them the cheapness of SSI in comparison to mitigation efforts seem to motivate there interest. The point is that this literature, which is aimed at policy-makers, takes SSI (as an example of SRM) seriously. It is clear that SSI meet the real world concern criteria from part 1.2.1 of this introduction.

The second condition to meet is that of SSI raising issues of ethical concern. There are many ways in which SSI raise ethical causes of concern (Preston, 2013). The case in favour of SSI meeting the ethical causes of concern criteria becomes clearer in chapter 2 when specific ethical concerns about SSI are explored. For the moment it is sufficient to draw attention to the fact that SSI is capable of inflicting harms which violate basic rights.[41] SSI gives the agents who control it the power to effect the background conditions of live, used poorly or maliciously an agent could threaten the background conditions of life for many, by doing so violate many basic rights, at an extreme, the agent who controls SSI holds the power to inflict a mass extinction upon the world. Yet even if SSI is done well it still risks violating basic rights, for example when sulphur leaves the stratosphere it can enter earth systems via acid rain, which can be fatal, hence it can contribute to the violating of basic rights.

SSI clearly meets both the conditions set out early in this chapter of being prominent in real world discussion and raising ethical concerns. It is therefore an appropriate type of geoengineering technology for this thesis to focus on. Therefore the problems of moral corruption, moral hazard and intergenerational inclusion raised in the context of SSI. It may be the case that the ways to address these problems in the case of SSI are applicable to other geoengineering technologies, yet this possibility is not explored in the thesis, given the relvance of SSI the thesis solely focuses on it.

1.3.2 Narrowing the scope of the thesis: clarifying the level of governance

Any project on geoengineering governance should provide a degree of clarity about the level of governance which it will focus on. There are several domains of governance which could seek to control and legislate on geoengineering practices within their domain. This could range from the regional level to the national level to international blocs right up to the level of global

[41] Earlier in this chapter the case was made that there are a plurality of reason for thinking that geoengineering could raise ethical concerns, and that violation of basic rights is one of them.

governance. There is no 'right' level of governance to be found, as was explained earlier all of these levels of governance are of interest and have a role to play. Nonetheless one needs to be aware of the level of governance which they are focusing on. It seems to me that the level of global governance is sufficiently interesting to be studied, but that does not mean that it is more important or interesting than other levels of governance. I will therefore explain why this thesis chooses the level of global governance for its focus.

The main reason why I choose to focus on the global level of governance is that it seems particularly well placed to address certain ethical problems which I argue are very serious. These are the moral hazard problem, moral corruption and how to include future generations. This is due to the nature of climate change being a global problem and consequently that geoengineering has global implications. Hence, many of our ethical concerns with geoengineering also have a global scope. Consider the issue of the moral hazard. The idea that the mitigation behaviour of agents might be adversely affected by SSI is something which is true of agents in any part of the world. Therefore, even if an international bloc such as the European Union (EU) had perfect principles of governance in place to address the moral hazard of SSI, it would not be able to address the problem in its entirety, due to there being relevant agents acting outside the EU. As explained this is not to claim that the level of global governance is the most interesting or appropriate to govern SSI just that it is sufficiently interesting to think about.

The appeal of focusing on the global level is compounded when we consider the different reasons people have to be interested in the climate. As already explained, the climate is a background condition of life. Consequently, everyone has a strong interest in the climate. Given the strength of this interest, it is difficult to conceive how a governance arrangement for SSI could be considered just if it excluded the representation of the strong interests of agents from all parts of the world. This is to say that SSI governance has to happen at the global level, just that the global level if sufficiently interesting to be the focus of this project[42]

[42] Although the analysis of these problems is carried out with the global level in mind, it may well be the case that much of the analysis is also applicable to other levels of governance.

1.3.3 A comment on method

Here I provide a comment on the methodological commitments underlying the thesis. There is no unifying methodology to this thesis in part due to the variety of tasks it performs. As mentioned, this thesis engages with distinct but mutually advantageous tasks of conceptual clarification and normative argument, as well as considering how a normative principle could be institutionalised. Therefore, I elaborate on the methodological commitments which are adopted in the normative components of this thesis.[43] A reason for highlighting one's methodological commitments is that they are indicative of what expectations we can have about a project. Thus by outlining my methodological commitments, I seek to provide a sense of the theoretical space in which the project is operating.

A number of different distinctions have been drawn in the recent literature on methodology in normative political theory, such as whether the project is one which assumes full or partial compliance (Rawls, 1999: 8; Valentini, 2012), whether it is realistic or utopian (Waldron, 1999, Williams, 2005, Cohen, 2008; Valentini, 2012: 4), or whether it is focused on transitional or end states (Rawls, 1999; Sen, 2006; Stemplowska, 2008). Hence I shall comment on what I understand to be the most relevant of my methodological commitments, which is how I understand the factual constraints under which normative political theory should operate; specifically, how I understand political constraints. My understanding of constraints is informed by Gilabert and Lawford-Smith (2012), who usefully identify two types of constraints: hard and soft. The former are constraints which will always be present, whilst the latter will not necessarily be forever present.

The distinction between hard constraints and soft constraints can be understood by considering two features: whether the constraint is malleable and whether it is understood in probabilistic terms. Phenomena which can be understood as possessing both these features can be understood as a soft constraints (Gilabert and Lawford-Smith, 2012: 814). Social, political, and economic institutions are malleable; this can be seen by observing the history of these institutions in which we can see radical changes from ancient societies to the present day. In contrast, laws of science are taken as hard constraints; gravity is not malleable in the sense that political institutions are.

[43] Specifically chapters 3, 5 and 6.

I do not consider this thesis to be bound by soft political constraints, such as the type of political institutions which currently exist or which we think are likely to exist in the future. Instead, the thesis is interested in understanding the potential for principles to address specific ethically concerning features of SSI governance. Chapter 6 also considers how one of these principles could be realised in an institution. This results in a novel political institution being proposed in order for future generations to be represented in SSI governance, as opposed to considering incremental changes to current institutions in the hope of representing future generations. This type of question is a legitimate pursuit in the realm of political theory, and is not contingent upon facts about prevailing or likely political institutions.

This type of methodological approach is also not uncommon when considering questions of climate justice. In particular, I understand it as being consistent with the view that Gardiner adopts when he grounds his argument for a global constitutional convention in political reality (Gardiner, 2014: 306).[44] Gardiner explains that by taking political reality seriously, we understand the gravity of the problem at hand, and the failure of current institutions to address it (Gardiner: 2014: 306). In this case, we understand the seriousness of the ethical concerns which are faced by SSI governance, their relation with the global institutional failure to address climate change, and we are aware of the absence of meaningful action to address these ethical concerns by prevailing political institutions.[45] Secondly, taking political reality seriously means that the proposed principles are ones which are aware of the severity of the problems and are fit to address them. This second feature of political reality emphasises the importance of not being bound by soft political constraints, which otherwise may prevent appropriate principles and solutions being proposed.[46]

It is noteworthy that those who do take the soft constraints of existing political institutions seriously can criticise those of us who do not as being unrealistic. To defend the position adopted in this thesis, it is not obvious that those who adopt soft political constraints have a monopoly on the idea of being realistic when it comes to considering ethical problems which institutions face. In particular, existing climate governance institutions have a 30 year history of failure to tackle the problem of climate change. If one wished to adopt these failed institutions as soft constraints, then one would require admirable optimism to think that such

[44] An argument which is explained in greater detail in chapter 6 of this thesis.
[45] Chapter 2 which is about framing is of significant relevance, for it is that chapter which seek to find the appropriate frame for problem at hand.
[46] This is clear when considering chapter 6 of this thesis, the ideas proposed such as a second chamber take seriously the need to represent future generations, but do not take soft political constraints seriously.

institutions could ethically govern SSI. Swift and White observe that we do not want political theory to be held hostage to public opinion; I would amend this statement so it reads that we do not want political theory to be held hostage to public opinion or *inadequate political institutions* (Swift and White, 2009: 63). And by not taking soft constraints seriously, we reduce the risk of political theory being held hostage by inadequate institutions.

1.4 Conclusion

In this introduction I started by outlining the problem of climate change and how geoengineering may be a tempting solution for addressing it, given our collective failure to tackle climate change or prevent it from happening. This raised the question of what geoengineering is. I used the Royal Society report to explain it, and highlighted some causes of concern with the Royal Society report raised by Gardiner and Jamieson. Once it was clear what geoengineering is, I then tried to explain why we would want it to be governed. This emphasised the ethical causes for concern with geoengineering. It is also clear that different types of geoengineering raise different questions of governance and therefore I decided to focus on the case of SSI. This is due to its prominence in real-world discourse and the strength of the ethical concerns we have with it. The introduction then outlined some principles of governance which have been proposed in the past by drawing on the work of Morrow. This was followed by explaining the aim of the thesis, which is to consider the ethical problems of moral corruption and moral hazard and of including future generations and to see whether the principles of governance which have been proposed to date are actually fit for purpose.

2 Context and Frameworks

The previous chapter has motivated an interest in the principles of governance for SSI. The purpose of this chapter is to give the reader the tools they need to understand the arguments presented in this thesis. To achieve this, two tasks will be performed in this chapter. One of the tasks is that of familiarising the reader with two frames which are used in all subsequent chapters of the thesis, those of the risk-risk trade-off, and the perfect moral storm. The other task of this chapter is to provide the reader with a sense of the ethical concerns which can be raised about SSI. This is necessary for a proper understanding of the risk-risk trade off frame. It also serves a secondary purpose of contextualising the core problems of moral hazard, moral corruption and intergenerational inclusion, which sit amongst a rich tapestry of concerns about SSI.

2.1 A brief overview of a brief history of ethical concerns with geoengineering

The goal of this subsection is to give the reader a sense of the ethical terrain which exists in the SSI. To do this, I will draw the reader's attention to some of the key considerations in the literature on SSI ethics, this will build upon the causes of ethical concern with SSI which were presented in chapter 1.[47] This subsection is not a complete overview of the literature, there are a plethora of ethical concerns with SSI, nonetheless this overview will allow the reader to situate the ethical concerns which this thesis focuses in the context of a broad set of ethical concerns about SSI. Additionally, this sense of the ethical terrain should help the reader understand the following subsection on the risk-risk trade off frame of SSI.

Recall in chapter 1 the case for SSI being an issue of ethical concern. The argument pointed to the fact that there is a plurality of reasons for finding SSI ethically concerning. Particularly SSI raises the prospect of egregious harm, by being able to violate people's basic rights. Additionally, it opens up the prospect of an abdication of responsibility by the wealthy nations of the world. This chapter builds on that preliminary discussion of harm and responsibility by outlining four more specific sites of ethical debate about SSI: objections of hubris; concerns about intra and intergenerational justice; problems of framing; and questions about how SSI compares to other responses to climate change. I go on to show how these issues are often grounded in more basic concerns about harm and or responsibility.

Prior to 2009, there is a significant absence of literature on the ethics of geoengineering, at least by philosophers. Dale Jamieson appears to be the only philosopher to have written an article about the ethics of geoengineering (Jamieson, 1996). In this article, Jamieson proposes four conditions which any response to climate change, including geoengineering should meet. It must be technically feasible, have predicable consequences, produces a state which is preferable on socio-economic grounds, and does not violate well-founded ethical principles (Jamieson, 1996, 326). And from here we have the birth of philosophical literature on geoengineering ethics, which seems to be immediately forgotten about for the next decade.[48]

Although an awareness of ethical concerns did exist in some legal and scientific scholarship geoengineering (Bodansky, 1996; Keith, 2000; Crutzen 2006). Notably, Bodansky drew

[47] For other overviews of the literature on geoengineering ethics see Preston (2012, 2013, and 2016) or Callies (2019a:11-14).

[48] That is not to say that the article is forgotten about, but that the philosophical community did not really engage with Jamieson's proposals on the ethics of geoengineering for a decade after its publication.

attention to the features of geoengineering which could be politically and legally problematic such as its global impact yet uneven distribution of impacts as well (Bodansky, 1996). Moreover, an awareness of ethical concerns can be seen in some early scientific literature on geoengineering, for example Keith (2000) warns of the danger of the slippery sloop, technical fix arguments, and moral hazard. Additionally, in the article by Paul Crutzen, which crucially lifted the taboo on geoengineering research, the point is made that geoengineering does raise ethical issues (Crutzen, 2006: 217).[49] Despite these publications, there appears to be a lack of engagement with the ethics of SSI prior to 2009.

Yet from 2009 there was a shift in the engagement from philosophers with geoengineering, this shift at least correlates well with the Royal Society report being published, in which it explicitly states that "the greatest challenges to the successful deployment of geoengineering may be the social, ethical, legal and political issues associated with governance, rather than scientific and theoretical" (Shepherd, 2009 p xi). This brings what Christopher Preston terms the first wave of publication on geoengineering ethics (Preston, 2016: x).[50] Whilst I will not adopt the wave framing for this section,[51] it does correctly indicate that 2009 onwards is the beginning of a significant number of publications on geoengineering ethics.

One of the most intuitive concerns about SSI is that of the hubris objection (Jamieson, 1996; Hamilton, 2013a, 2013b; Hartman, 2018; Callies 2019a: chapter 2). In general, hubris objections are concerns about playing god, in context the fear is that performing SSI would be playing god. There are at least two ways that this objection can be understood, it could be a claim that the use of SSI requires a god like control over nature which humans should not have (Jamieson, 1996). Or there is the claim that humans simply cannot have the knowledge to understand the impact of using SSI and that this should at least have implications for how we approach the use of SSI (Hamilton, 2013a, 2013b).[52]

Second, it is widely acknowledged that SSI raises concerns of both intra and intergenerational justice. The literature on intragenerational justice is motivated by a concern about how different

[49] Although the article does not elucidate the ethical concerns, it merely draws attention to the fact that they do exist.
[50] For Preston the first wave is characterised by asking questions of geoengineering technologies on their own terms, which means in the case of SSI, the first wave of geoengineering ethics would solely ask questions of the benefit and burdens of SSI in isolation, not by comparing them to alternative climate outcomes. This is in contrast to the second wave of geoengineering ethics which is understood as publications on geoengineering ethics which are in compared to the harms faced by climate change (Preston, 2016: xii).
[51] The reason for not adopting Preseton' wave framing is that it is an unnecessarily complicated way of grouping arguments about SSI for a reader who is unfamiliar with the literature on SSI ethics.
[52] For a detailed account of the Hubris objection see Hartman (2018).

groups will be adversely affected by SSI. These concerns tend to be grounded in the fact that not all agents have the same adaptive capacity the adverse impacts of SSI, and that the distribution of these impacts is uneven. There is distinct intragenerational literature on how the adverse impacts of SSI would affect the vulnerable and how this raises tough theoretical questions about SSI, such as those of consent, legitimacy and justice (Hourdequin, 2012, 2016, 2018, 2019; Preston, 2012; Whyte 2012; Davies, 2013; Morrow *et al*, 2013; Svoboda, 2016, 2017; Carr and Preston, 2017; Calles 2019). The early texts on this subject tend to focus on identifying that where vulnerability exists and that such vulnerability points raises an issue of justice, or that it is a source of ethical concern (Hourdequin, 2012; Preston, 2012; Whyte 2012). Then more contemporary texts seem concerned about how we account for this vulnerability (Hourdequin, 2016, 2018, 2019; Carr and Preston, 2017; Smith 2018; Callies, 2019a). Such approaches tend to understand procedural justice as at least part of the cite to address this vulnerability.

There are also intergenerational concerns raised; these concerns tend to be about how SSI could wrong future generations. This ranges from concerns about the type of relationship that SSI would create between present and future generations (Smith, 2012, forthcoming), to how we could wrong future generations by presenting them with a future in which they have to make an undesirable choice about SSI (Gardiner, 2011; Ott, 2012). One of most prominent intergenerational concerns is that of the termination shock, which is a concern about how future generations could be adversely affected if they were to stop using SSI (Ott, 2011), this concern is considered in risk-risk trade off section of this chapter.

There does appear to be a difference in the development of intergenerational literature as oppose to the intragenerational literature. The intragenerational literature has recently focused on how to address issues through institutional design (Hourdequin, 2016, 2018, 2019; Preston and Carr, 2018; Smith 2018; Callies, 2019a). Whilst there does not appear to be any literature focused on designing SSI governance institutions to avoid wronging future generation.[53] At best there are suggestions about the importance of aggressive mitigation efforts to minimise the degree of the dilemma that future generations are placed in (Ott, 2012).

Thirdly, issues of framing are also amongst some of the first to be raised about SSI. These concerns tend to be of the same spirit, which is that certain framings of SSI mislead or obscure

[53] This absence of literature on how to address intergenerational issues serves as a strong motive for chapter 6 of this thesis.

important ethical considerations (Scott, 2011; Carr *et al*, 2011; Gardiner, 2013; Gardiner 2014, Morrow 2014; Fragnière and Gardiner, 2016; Heyward, 2016; Mclaren, 2016; Gardiner and Fragnière, 2018). The plan B / insurance framing of SSI are amongst the most heavily criticised frames of SSI; both of these frames make SSI sound unproblematic or even desirable (Scott, 2011; Gardiner 2013; Fragnière and Gardiner, 2016*).* The public good framing of SSI has also been subject to debate, as explained in chapter 1 this framing of geoengineering is made prominent by the first Oxford principle describing SSI as a public good. Since then, Gardiner has been critical of this framing on the grounds that geoengineering is not necessarily public or a good in the technical sense (Gardiner, 2013; 2014; Gardiner and Fragnière, 2018). Although the public good framing of the principles has also been defended on the grounds that geoengineering does meet the technical criteria for being a pubic good, which means that the use of it does not prevent someone else using it (non-rival), and that an agent cannot prevent others from benefitting from it (non-excludable) (Morrow, 2014). The theme of these articles is to observe how certain frames hide what is ethically at stake when considering SSI. Framing articles tend to be negative as well, in the sense that they tell the reader, do not frame geoengineering in terms X, as oppose to do frame it in terms Y.

More recently there have been a development of articles which seek to understand what should be done in light of alternatives to geoengineering, in fact this is the explicit theme of Preston's second edited volume on geoengineering ethics (Preston, 2016). For example, it is well established that SSI ought to be accompanied by aggressive mitigation if it were to be deployed (Baatz and Ott, 2016; Morrow, 2017). Yet a more radical position in the literature now exists, it accepts the idea that SSI ought to be accompanied by aggressive mitigation if it were to be deployed, but then goes on to state that we are likely to be in circumstances where we have a moral duty to in act this mixed response (Baard and Wikman-Svahn, 2016; Horton and Keith, 2016).

We should now have a sense of where some of the core concerns about SSI; objections on the grounds of hubris, concerns about intra and intergenerational justice, fears of it being framed poorly, and questions of how it compares to other climate change responses, these are all significant concerns with SSI. This sense of the ethical terrain is valuable insofar as it contextualises the concerns which this thesis focuses on.[54] To contextualise the concerns of

[54] As explained at the beginning of this subsection, this sketch of the ethical terrain should be consistent with the general causes of ethical concern which were presented in chapter 1, the ethical concerns expressed her can be understood as specific instantiations of the concerns about harm and responsibility in the case of SSI.

this thesis, the concerns of moral corruption, moral hazard and intergenerational inclusion are the type of concerns which could stand beside the concerns which have been sketched in this subsection if one wished to provide an account of the ethical problems with SSI. If one wanted a complete explanation of the ethical concerns associated with SSI, they would at least have to explore all the concerns mentioned in this subsection, we well as many others.

At the start of this subsection I expressed the view that these specific concerns are related to the causes of ethical concern identified in chapter 1; in some cases these are substantiations of the broad concerns expressed in chapter 1 about harm and responsibility. I shall now show why this is the case, by doing so I will show that the causes of ethical concern from chapter 1 and this subsection are complementary to each other. Consider the intra and intergenerational concerns with SSI. Many of these concerns point to ways in which vulnerability exists in the intra or intergenerational context (Gardiner, 2011; Hourdequin, 2012; Ott, 2012; Preston, 2012; Smith, 2012; Whyte 2012). The expression of the concern of vulnerability often draws both on considerations of harm and responsibility. For a vulnerability complaint to exist, there has to be an X which the agents are vulnerable to, this X is often a harm from SSI, such as changes to the African and Asian Monsoons. Moreover, there is often the observation that these vulnerable parties bear little or no responsibility for circumstances in which SSI would be used. The point about vulnerable parties lacking responsibility for the use of SSI draws attention to the perniciousness of the harms which would occur from SSI. Considerations of harm and responsibility seem to be common when articulating intra and intergenerational ethical concerns with SSI.

The same is true of framing, as established concerns about framing which have been expressed in the literature tend to point to how certain frames hide what is ethically important in the case of SSI. Given that harm and responsibility are important ideas it follows that framing concerns can be expressed when either of these points are hidden by a frame. For example, some of concerns about the plan B and the global public good framing is that they both divert our attention form the harms of SSI (Gardiner, 2013; Fragnière and Gardiner, 2016).

This is not to claim that all ethical objections to SSI will be grounded in harm or responsibility-based concerns, the concern of hubris might be an example of one which does not fit as neatly into a concern about responsibility or harm. The relationship between the general cases for SSI being an ethical cause of concern from chapter 1, to the specific formulations of ethical concerns which have been presented in this subsection should now be clear.

Now that we have a broad sense of the type of concerns which can be raised about SSI, we can explore the risk-risk trade off framing. This frame requires a sense of the risks from SSI, which is why we need to bear in mind the ethical concerns raised in the previous couple of pages.

2.2 The risk-risk trade-off frame.

As explained in the previous section, there are many framings of SSI which are problematic, such as the insurance, plan B and exceptionalism framings *(*Fragnière and Gardiner, 2016; Heyward, 2016). This is indicative of a trend in the literature on geoengineering and frames, which is that this literature is negative: it argues against rather than in favour of particular frames. Yet there are also framings which are helpful for capturing what is at stake. Prominently, there is the risk-risk trade-offs framing of SSI, which appears to be implicit in many discussions about SSI and explicit in the work of Reynolds and Fleurke (2013), Baatz (2016) and McKinnon (2018). The premise behind this frame is that any position one could adopt on SSI, ranging from a moratorium to research now, is one which comes with some risks. Therefore when adopting any position about SSI governance, one is trading risks off against each other.[55] The FCEA report identifies three locations of risks which this framing captures *(*Chhetri *et al*, 2018: 4). The locations are the risks from researching or not researching SSI, the risks of deploying or not deploying SSI, and the risks of governing SSI. In this subsection I shall give an account of these risks to clarify the framing, which will then be used in the thesis to contextualise discussion. There is an additional benefit of exploring this frame, which is that it should give a clearer sense of specific complaints which can be raised about SSI.[56]

The first risk which is often raised when thinking of researching SSI is the moral hazard concern. As explained earlier, the intuition behind this concern is that policymakers or the public may reduce their mitigation efforts in response to SSI research (Shepard *et al* 2009; Hale, 2012; Lin 2013; Morrow; 2014). An additional risk is that of technology lock-in, whereby the fact that a technology like SSI is being researched starts creating pressure for that technology to be deployed even if there is a lack of consensus about whether the deployment is desirable (Preston, 2013; Hulme, 2014; Callies, 2019a: chapter 2; McKinnon, 2019). There is also a risk that this technology may end up being harnessed by special interest groups, for example the military (Gardiner, 2013: 29; Long and Scott; 2013).[57] It may also increase the

[55] In this respect there is a stylistic similarity between the risk-risk trade off frame and the works from Preston's edited volume (2016), where a theme of the volume is that the chapters are comparative to climate change scenarios.

[56] The intention of this subsection is to use the risks associated with SSI which were not considered in the previous subsection so that the reader gets a broader sense of the ethical concerns which exist about SSI.

[57] The reason behind this concern looks more compelling if one recalls previous military endeavours into weather invention, as can be seen in Fleming (2010), a prominent example being how the United Sates engaged with weather modification actives in the Vietnam War (Fleming, 2010 179-182)

possibility of rogue deployment, for more agents will have knowledge about how SSI works (Gardiner, 2013: 29).[58]

Yet as the risk-risk trade-offs framing implies, there are also risks which exist if SSI research does not happen. They are risks which occur from an absence of knowledge about SSI, these risks appear to range from we could perform SSI in a suboptimal manner, to we would use it dangerously, or that we would not be able to use it all. The suboptimal complaint is that we would not be aware of certain direct benefits from performing SSI (Keith 2013, 2017: Horton and Keith 2016: MacMartin, *et al* 2018) and that we would lack the knowledge of SSI to integrate into policy to pursue co-benefits from its use SSI (Buck, 2012: Stelzer and Schuppert 2016). One of the most appealing co-benefits is the idea of SSI being able to help alleviate human development problems. It could be beneficial to crop growth and subsequently improve food security (Robock, 2016), it could allow people in poverty to use dirty fuels which may be banned in an aggressive mitigation scenario, and by doing so help alleviate energy poverty (Buck, 2012). Yet to fully understand the potential co-benefits of SSI, proper research into the impacts from SSI is required, if we are going to obtain knowledge of this and the other benefits associated with SSI. The dangerous complaint goes that if we lack the knowledge of how to use SSI safely, we are more likely to be in a situation where the use of SSI results in the imposition of dangerous risks, therefore research is required to reduce this risk (Keith *et al*, 2010). There is also the concern that if we do not research how to deploy SSI, we would be unable to deploy it responsibly, it is estimated that it would take 20 years' worth of research to know how to do this (Chhetri *et al*, 2018: 4).

The risk-risk trade off also frames a way of thinking about the deployment of SSI. Some of these risks have already been explained in the introduction, such as the risk of undesirable climate impacts such as acid rain, and the fact that they are not equally distributed for example, an adverse impact upon the Asian monsoon (Kravitz *et al*, 2009).[59] Moreover, just as the distribution of impacts is uneven across space, they are also uneven across time. It is possible that future generations will face worse climatic impacts than the generation which deploys SSI (Gardiner, 2011: chapter 10). Additionally, there is the risk of termination shock: the fear that if SSI were deployed, then any generation which discontinued the deployment of SSI would face rapid climate change. The reason for this is that SSI only hides the effects of some

[58] Rogue deployment could occur from non-state actors or a set of states. Gardiner terms this a coalition of the willing (Gardiner, 2013:29).

[59] This concern has a relation to Hamilton's hubris objection, namely that we cannot understand earth systems well enough to be able to predict all the risks.

greenhouse gas emissions. Thus if SSI stops, it stops hiding the impact of greenhouse emissions and the global temperature adjusts to levels it would be independent of SSI *(*Brovkin *et al*., 2009; Irvine *et al*., 2012; Jones *et al*., 2013; Llanillo *et al*., 2010; Matthews & Caldeira, 2007; McCusker *et al*., 2012, 2014). The reason why this is feared is that the speed of this adjustment is meant to be fast, much faster than following a climate trajectory independent of SSI (Matthews & Caldeira, 2007; Llanillo *et al*., 2010; Irvine *et al*.,2012; McCusker *et al*., 2014). This faster rate of change increases the amount of harm done by CC, due to there being less time to adapt to these harms (McCormack *et al*., 2016; Trisos *et al*., 2018).[60] This does appear to be one of more discussed risks about the deployment of SSI.

The risk of not deploying SSI is the risk of facing a world in which the temperature has increased, if we were to the follow our current projected pathway, by 3 degrees (Rogelj, *et al*, 2016). *A*s was also explained in the introduction, we have reasons to care about the climate, a prominent one being the fact that we are dependent upon it. This risk appears to motivate much of the interest in SSI research and development, so that we could deploy SSI to reduce the impact of climate change. This logic can be seen in a significant number of recent writings on SSI (Crutzen, 2006; Barrett, 2008; Michaelson, 2013; Reynolds and Fleurke, 2013; MacMartin *et al*, 2014; Moellendorf, 2014; Horton and Keith 2016; Callies, 2019a; Svoboda *et al*, 2019). Of course, the severity of these risks is not set in stone; depending on our actions now, we could face a world of significantly more or less than 3 degrees of warming, and this would of course alter the risk of not deploying SSI. For example, no one appears to advocate for deployment at this point in time, they tend to advocate for the use of SSI at the point in time when either SSI is part of an optimal climate response or when the risks from climate change are simply too great (Reynolds and Fleurke; 2013; Moellendorf 2014; Horton and Keith 2016).[61]

[60] Although are also those who hold that the termination shock fear is overemphasised, Parker and Irvine (2018) think that the termination shock is simple to address, they do not think that political actors would choose to behave in such a way that the termination shock occurs, so all that is needed is to ensure that this appropriate protection of the infrastructure which deploys SSI, so that it cannot be damaged. This position is vividly critiqued on the grounds that it depends on 'heroically optimistically assumptions' about the motivations of political actors (McKinnon, 2019). Secondly there is Florian Rabitz (2019) who argues that the termination shock carries too much weight in discourse about SSI, and that it is no more problematic than risk which exist in nuclear non-proliferation agreements, are response to pandemic threats or global financial markets. All of which, if they were to fail would have devastating consequences.

[61] For those who are tempted by the idea of deploying SSI once climate change is bad enough, they would do well to heed the warning of Jamieson, who aptly observes that one person's climate emergency is another's bad day (Jamieson, 2014: 221). The point is that unless there is an agreement on what climate change being bad enough is then this attitude is going to result in significant disagreements.

There are also risk-risk trade-offs faced in the governance of SSI. For example, consider a very well-regulated governance regime which is designed to meet important normative criteria such as that of procedural justice, which requires that all relevant agents are appropriately included in the decision making process (Hourdequin, 2018, 2019; Callies, 2019a: chapter 6). Such a regime may take a long time to make decisions and effectively constrain our ability to respond to SSI at a meaningful speed. Vice versa, an SSI governance regime which is very fast at making decisions may fail to meet the standards of procedural justice, due to it restricting the set of agents who are able to exercise power in the decision making process. This is where the risk of substantial injustices to vulnerable groups exist, there is the risk that such a governance regime would fail to include future generations in some form and other marginalised groups (Hourdequin, 2012, 2016, 2018, 2019; Smith, 2012, 2018; Preston, 2012; Whyte, 2012).

Additionally, there is the risk of poor governance creating the conditions which facilitate the rogue use of SSI (Keith *et al*, 2010). The rogue use of SSI could occur for a variety of reasons, military and economic reasons have both been raised in the literature (Gardiner, 2013; Wagner and Weitzman, 2015). The economic reason is particularly concerning in its simplicity. As has been established, the cost of SSI could be in orders of magnitude less than that of the cost of the burdens of climate change to a single state. For example, the cost of deploying SSI is comparable to the cost of a Hollywood blockbuster (Keith, 2013: IX), therefore, states face a financial incentive to perform SSI. This creates the free-driver problem, whereby multiple states are incentivised to go rogue and pursue SSI policy, independently of the actions of other states (Wagner and Weitzman: 2015; 99-101). A well-regulated and enforced SSI governance regime should be capable of preventing the rogue use of SSI.

Earlier in this chapter, we considered the criticism of certain framings used in the SSI literature. A common theme of these criticisms is that problematic frames obscure what is ethically at stake by adopting inappropriate language, such as that of emergency or plan B. It seems that a similar concern can be raised about the risk-risk trade off frame; it could invite the reader to engage in an exercise of trading off benefits and burdens of climate change against those of SSI, in a problematic way.

To see the problematic way that comparing risks can occur, we need to consider some observations by Gardiner regarding cost-benefit analysis, particularly how the term can be used to refer to different techniques. The first technique is that of a pros and cons analysis; this is simply where the positives and negatives of a situation are identified and analysed, and this

technique in some form seems unavoidable when making ethical judgements (Gardiner, 2016: 77). Secondly, there is net benefit analysis; this is where the controversy starts. Net benefit analysis assumes two points: firstly that the benefits and costs are comparable, secondly that we should aim for that situation whereby the sum of the benefits is as high as possible (Gardiner, 2016: 77). This second assumption is particularly concerning in our case, as we are not indifferent between the distribution of benefits and burdens which occur from this risk-risk trade off. Despite the limitations of net benefit analysis, it is common place in many academic and public policy locations (Griffen; 1998; Pearce, 1998; Held *et al*, 2016) Given the prevalence of this net-benefit approach, I fear the risk-risk trade off frame could mistakenly invite a net-benefit analysis, which hides the ethical issues at stake when making a judgement about a position to adopt on the risk-risk trade off spectrum. The solution which this thesis adopts for this problem is to also use the perfect moral storm framing, which puts ethical considerations at the centre of our analysis; the way this works will be explained in the following section.

The risk-risk trade off frame has an additional benefit, which is that it provides a way to think about criticisms of ways to engage with SSI. Consider the governance spectrum which was introduced in chapter 1: one could argue in favour of a position anywhere on the spectrum, from a new global governance institution, to SSI governance based on norms, or a moratorium on governance entirely. The risk-risk trade off frame illustrates the point that none of these points on the spectrum of governance are risk free. They will all require trading off certain risks against each other. This is highly important, since it shows that the fact that a position on the governance spectrum comes with risks is not a reason to reject that position. For the presence of some risks is unavoidable when taking a position on the type and form of SSI governance. Therefore, when forming judgements about SSI we know that risk is in part a price of having that judgement. The risk-risk trade off frame effectively draws attention to this idea.

The risk-risk trade off frame can be used to show how strong arguments regarding geoengineering are underdeveloped. Nightingale and Cairns (2014), have a powerful and useful paper in which they systematically draw attention to the security risks which researching SSI creates. One of their key arguments is that SSI will end up under the control of the military not scientists (Nightingale and Cairns, 2014: 5). The conclusions of their paper is that there should be a moratorium on SSI research due to the unacceptable security risks which it poses. This is persuasive, but the risk-risk trade off frame shows us that their argument is not yet complete; the question which the risk-risk trade off frame invites us to ask is ‘would these risks still exist

if SSI research was banned'? Plausibly, it is not clear that banning the research of SSI would prevent or even reduce the likelihood of militaries researching it. The risk-risk trade off frame invites important critical questions to be asked regarding their paper, the omission of these questions makes their argument appear somewhat incomplete.

One may accept the point about the risk-risk trade off having virtues in comparison to cost benefit analysis, but wonder whether it is desirable in comparison to other possible frames which we could adopt. As explained earlier in the chapter, most of the published work on geoengineering and frames is done in the negative style of criticising frames, as opposed to arguing in favour of a frame. Here I will provide an example of how the risk-risk trade off frame compares favourably to an alternative possible frame.[62] The alternative frame being a precautionary approach to geoengineering.[63]

When thinking of the precautionary principle, it is important to be clear that this principle has been understood and used in very different ways to generate different conclusions (Elliot, 2010; Hartzell-Nichols, 2012; Reynolds and Fleurke, 2013; Tedsen and Homann, 2013; Morrow, 2014). Therefore one cannot claim to analyse what the precautionary principle will prescribe in this case; instead I need to specify how a particular understanding of the precautionary principle could function. Here I will adopt Lauren Hartzell-Nichols' understanding of precautionary principle, which I think captures it well. Hartzell-Nichols defends a catastrophic precautionary principle when making judgements about geoengineering. This principle states that 'Appropriate precautionary measures should be taken against threats of catastrophe' (Hartzell-Nichols, 2012: 160).

To demonstrate the different impacts of these frames, let's consider what these frames tell us about the prospect of SSI being researched. Hartzell-Nichols comments on how the catastrophic precautionary principle would endorse research into SSI, as long as such research does not pose a catastrophic risk (Hartzell-Nichols, 2012: 167-168).[64] As Hartzell-Nichols argues, small-scale research does not appear to generate catastrophic risks; therefore the

[62] The risk-risk trade off frame is compared favourable to the insurance frame in part 4.2 of the thesis when considering why the classic insurance understanding of the moral hazard problem is inappropriate when the moral hazard objection is raised in the case of SSI.

[63] It is important to note that the precautionary frame is not the only way of understanding the precautionary approach, it is commonly understood as a principle as well. There is perhaps certain oddness in describing the precautionary principle as a frame, I suggest that the fact that precautionary principle is being used as a frame despite it not being the most natural use of the precautionary principle should be understood as indicative of the lack of positive frames which have been proposed to consider geoengineering.

[64] The precautionary principles of Elliot (2010), Reynolds and Fleurke (2013), and Morrow (2014), reach comparable conclusions of endorsing limited research into geoengineering.

catastrophic precautionary principle would endorse research into SSI as a guard against the catastrophic risks of climate change. Moreover further research into SSI would be endorsed insofar as there is no reason to think that such research imposes a catastrophic risk.[65] In contrast, the risk-risk framing should bring all relevant risks to the surface regardless of their catastrophic nature; the insights in this case from the risk-risk trade-off frame have already been explained in this section.

The precautionary principle comes with a strong normative commitment to avoiding catastrophe, and that normative commitment is informative of much of the analysis which comes out of that principle. Here we can see a distinct difference from the risk-risk trade off frame, which comes with no such normative commitments. This points to a distinct advantage of the risk-risk trade-off for this project, since the background normative commitments of this project have already been outlined in Chapter 1. Thus the task of this chapter is to provide a frame which will highlight the relevant features of geoengineering. Once we are aware of these features of geoengineering we can consider them in light of our normative commitments. The ability of the risk-risk trade off to provide a lucid description of a situation without drawing upon strong normative commitments is a strong reason to adopt this frame for this project.

In summary, the risk-risk trade-off frame is helpful because it provides a way of thinking about the nature of the ethical challenges involved in SSI governance. It is a fact that any position on SSI governance, ranging from a moratorium, to extensive governance of all stages of SSI at the level of a global institution, comes with risks. Any position adopted in this thesis is one which will be making judgements and trading off some of the risks against each other.

[65] The termination shock can be understood as a catastrophic risk, as things stand it appears that the catastrophic precautionary principle would reject research which creates conditions for the termination shock to happen, the truth of this is not contingent upon how unlikely researchers think the termination shock is to happen.

2.3 The perfect moral storm

This section shall provide an explanation of Gardiner's perfect moral storm (Gardiner, 2011). The reason for using the perfect moral storm is that it offers a useful framework for understanding the problem of climate change, and the ethical problems which arise from SSI, including those of moral corruption, the moral hazard, and the intergenerational problem of inclusion. Importantly, this should not be seen as a competing framework to the risk-risk trade off, since they focus on different issues. Whilst the risk-risk trade off frame presents an understanding of how judgements about SSI are made, the perfect moral storm presents an account of the background conditions against which climate change policy occurs. The complementary nature of these frames is explored at the end of this subsection, once the perfect moral storm has been explained.

Gardiner's project in a Perfect Moral Storm is to understand the nature of the problem of climate change.[66] He tries to offer an uncontroversial characterisation of this problem, by framing it in ways that all morally serious people should be able to accept. The purpose of the book is to provide a prerequisite for identifying desirable solutions to climate change. We cannot find appropriate moral theories if we do not understand the nature of the problem. Therefore, the Perfect Moral Storm is a vital tool for being able to develop appropriate moral theories to combat climate change. The method he chooses for this framing is centred on the notion of storms. The choice of the storm frame is due to storms being bad, and the fact that multiple storms can converge to make a situation much worse. Gardineridentifies the convergences of three different storms which threaten our ability to ethically respond to climate change:the global, intergenerational and theoretical storm (Gardiner, 2011: 11). He acknowledges that there are possibly more storms, such as an ecological storm, but the tools required to identify them require more controversial assumptions, therefore he avoids engaging with them.[67]

The first part of the book focuses on the global storm, which describes the global environmental problem with a sole focus on the present generation. To understand the nature of the climate change problem, Gardiner considers the tragedy of the commons framing; such a frame has a

[66] Technically Gardiner understands the storms as understanding the global environmental problem of which climate change is a core case. Given the focus of this thesis I shall use the perfect moral storm as a way of understanding climate change, not the global environmental problem.

[67] In later writing Gardiner does give more detail about the ecological storm (Gardiner, 2016), although he is still yet to provide a detailed account of this storm.

prima facie appeal to explain the structure of the climate change problem and is notably adopted in the work of Posner and Weisbach (2010). The tragedy of the commons is a situation where individually rational behaviour undermines the possibility of reaching an outcome which is collectively rational (Hardin, 1968). In the case of climate change, the logic suggests that it is collectively rational for states to limit the total amount of global emissions to avoid harms from climate change, yet no country wants to the bear the burden of cutting their own emissions, they would rather free ride (Gardiner, 2011: 28). The tragedy of the commons is a tricky situation, but one where is there much literature; a common solution is that of having an institution with coercive power to make it individually rational for agents to do what is collectively rational (Vogler, 1995; Hall, 1998).[68] This means that the climate change problem could be addressed, if one were to have the appropriate global governance institution (Gardiner, 2011: 29).

Gardiner argues that addressing climate change is not as simple as the tragedy of the commons understanding would have us believe. Gardiner offers three reasons that highlight how the tragedy of the commons understanding of the climate change problem fails to capture the role of scientific uncertainty, deep roots and skewed vulnerabilities (Gardiner, 2011: 29-32). The concern of scientific uncertainty is that it is not clear how each nation will be affected by climate change, and it may be case that nations are not convinced they are better off in a world in which climate change it addressed.[69] The case for such uncertainty is compounded when we consider the deep roots problem. Here Gardiner observes that those agents with a vested interested in the continuation of a carbon intensive system wield economic and political power (Gardiner, 2011; 30-31).[70] Hence you have wealthy agents funding the research and publications of papers which deny that climate change is a problem, or even deny the existence of climate change (Dunlap and Jacques, 2013; Elsasser and Dunlap, 2013). Thirdly, Gardiner observes that there is the problem of skewed vulnerabilities. This is the problem that those countries which have contributed the most to climate change are those which are the least likely to be affected by it in the short term. The reason for this is twofold: due to their wealth they are better placed to adapt to the effects of climate change, and geographically they tend to be located in temperate zones. That means those states which are most at risk from climate change

[68] For example, an institution could punish those who overuse their share of a common to the point at which the optimal use of that commons for individual is the same as what is rational for the collective.
[69] This could be in the sense of comparatively better off than other states under climate change.
[70] The concern of deep roots is also used to observe that responding to climate change may require changes in the way people live their lives, the extent of this change will in part depend on the availability of renewable energy (Gardiner, 2011: 31).

are those who contributed the least to it (Gardiner, 2011:119). Gardiner also makes the helpful point that countries do not participate as equals in the international system, some are much more powerful than others and history is littered with powerful countries abusing their power. This raises the concern that powerful countries will engage in exploitative practices when responding to climate change, this has been labelled as compound injustice (Shue 1996; Gardiner 2011). Combined, these concerns show why the problem of climate change is worse than the tragedy of the commons approach would have us believe, as it is not the case that all states are equally affected and equally able to respond to climate change.

The rejection of the tragedy of the commons model brings many useful lessons. One which bears particular relevance to this thesis concerns the danger of skewed vulnerabilities and compound injustice. The lesson of skewed vulnerabilities is that some agents will be particularly vulnerable to decisions made about SSI. This is a product of geography, and lacking the means to adapt to the effects of SSI. Moreover, SSI governance could contribute to compound injustice by perpetuating current unjust global power structures.

The second storm is the intergenerational storm. One would be right in already feeling depressed about the global storm, yet the problems posed by the intergenerational present an even greater challenge. When setting up the intergenerational storm, Gardiner reminds the reader that climate change is resilient; carbon emissions will stay around for much longer than present generation, 20–60% for at least 1,000 years, therefore we cannot prevent the effects of climate change purely from mitigation efforts (Gardiner, 2011: 33).[71] Moreover, Gardiner reminds the reader that the total increase in temperature from the present generation's emission will not be felt until these emissions have had their life in the atmosphere. This means that the full impact of our emissions is not experienced by us, but by future generations. Gardiner then asks the reader to assume that states do not represent the interests of the citizens significantly far into the future, a plausible assumption (Gardiner, 2011: 35).

Gardiner then considers the question of how succeeding generations would behave in this situation.[72] He thinks that most generations would rationally prefer the outcome when there is sufficient mitigation, in order to avoid harms from climate change.[73] Yet when each generation has the power to mitigate, they prefer not to mitigate. The logic being that each generation

[71] This is assuming an absence of scalable geoengineering technologies.
[72] For the purposes of the analysis he asks us to imagine that generations do not overlap (Gardiner, 2011: 170).
[73] The reason Gardiner say most is because the first generation is not is not affected by the behaviour of any other generation, so does not have the same rational interest in other generations mitigating.

receives the benefits from its emissions, but burdens are borne by future generations.[74] Gardiner names this the pure intergenerational problem. Solving this problem is even harder than the conflict in the global storm, since with the global storm the question seems to be how you can design an institution which can get parties to behave in the way that they should do to each other. Yet in the intergenerational storm you cannot even get the relevant parties to partake in a coercive institution at the same time (Gardiner, 2011: 37). This raises the possibility of what Gardiner terms the tyranny of the contemporary, whereby each generation pursues its self-interest to the detriment of future generations. A reason why the tyranny of the contemporary is particularly nasty is that when a generation fails to mitigate, the severity of climate change experienced by future generations will increase. Given that we can have grounds to fear that each generation will fail to sufficiently mitigate the situation, it appears that the severity is going to compound over time.[75] The pure intergenerational problem also shows the inability of the tragedy of the commons system to understand the climate change problem, which instead appears to assume that a state represents the interests of all its future citizens. The usefulness of this for the thesis is that the intergenerational storm draws attention to the danger of the pure intergenerational problem and the danger of the tyranny of the contemporary in climate policy decision making.

The theoretical storm reflects the failure of moral and political theory to cope with challenges raised by climate change: in this list are "intergenerational equity, international justice, scientific uncertainty, persons whose existence and preferences are contingent on the choices we make, and the human relationship to animals and the rest of nature" (Gardiner, 2011: 213-214). Yet political theory is not alone in having theoretical limitations when it comes to engaging with climate change. Take cost benefit analysis, a tool which seems very poorly equipped to deal with analysing climate change despite being used to do exactly that, notably by William Nordhaus (2009), and Nicholas Stern (2007). The inadequacy of cost benefit analysis stems from the fact that it requires questionable moral judgements to work in the case of climate change, ranging from choosing a discount rate, to putting a monetary value on many components of the analysis, including statistical lives and environmental quality (Broome

[74] Gardiner provides a technical analysis of why this situation is worse than an intergenerational prisoners dilemma account of the problem would have us believe., I shall not explain that here, since it does not add much to the analysis of this thesis (Gardiner, 2011; chapter 5)

[75] Gardiner does provide an analysis of how this can result in a tragic future where severely burdened generations may have to emit if it is there best form of generating the energy and wealth required to adapt to the harms of climate change, but by doing so poses an even greater burden to the next generation (Gardiner, 2011: 39-41). And this could happen indefinitely, every generation unnecessary bears a huge burden and passes an even greater one onto the next generation.

1992, 2012; Gardiner, 2016). One of the main points that Gardiner wants to make is that many of our key theoretical tools in society are inadequate for dealing with the climate problem. Therefore, even if policy makers did want to overcome the global and intergenerational storm, we would struggle to provide them with the appropriate theoretical tools for action. Despite this, Gardiner thinks we can deal with the climate crisis in the short to medium term, by using techniques such as identifying cases of moral failure.

With this in mind, we should recognise that the perfect moral storm is a helpful account of the current situation in which climate change policy occurs. It proves a vivid rejection of the tragedy of commons framing of the climate change problem, which one might otherwise be tempted to adopt. By doing this, it draws attention to important considerations such as skewed vulnerabilities, compound injustice and the pure intergenerational problem. In the subsequent chapters, I will draw attention to how my arguments relate to these parts of the storm. For example, the following chapter on moral corruption is a product of the perfect moral storm as a whole. Additionally, chapter 6, which seeks to address the problem of intergenerational inclusion, is acting at the intersection between the intergenerational and theoretical storm. Broadly speaking, the perfect moral storm provides an account of the conditions against which our discussion of SSI is happening and why we might find it particularly tricky to address the ethical problems associated with SSI.

2.4 Combining the risk-risk trade-off and perfect moral storm frames

This chapter has introduced two different frames which will be used throughout this thesis. I shall provide a brief comment on the relationship between these two frames. The perfect moral storm provides a sophisticated account of the situation against which climate change policy occurs. Specifically, this account shows why it may be so hard for the current generation to make just responses to climate change. Given that the perfect moral storm can be understood as an important background condition against which judgments about SSI are made, the risk risk-trade off frame is a useful account of the type of judgments which one must make when choosing a stance on SSI policy. Trading off these risks may seem daunting in the best of circumstances, yet the perfect moral storm shows that we do not make that judgement in the best of circumstances, instead we make it in thoroughly challenging circumstances. The prefect moral storm then gives us reason to suspect that we will struggle to identify a just position when making these risk-risk trade-offs. A benefit of adopting the perfect moral storm framework to accompany the risk-risk trade of framework is that the perfect moral storm helps mitigate the concern that ethical concerns could get lost in a net benefit analysis approach to the risk-risk trade off frame. This is due to the perfect moral storm effectively highlighting the ethical challenges that responses to climate change face.

Earlier in this chapter, I argued that the risk-risk trade-off frame alone would invite ethically suspect action. Hence the need for a moral framework such as the perfect moral storm to be combined with the risk-risk trade off frame to mitigate this danger. Now I shall consider a different question and provide a reason for thinking that it is beneficial to have the risk-risk trade-off frame in addition to the perfect moral storm frame. By doing so I illustrate how each frame benefits from being combined with the other. The perfect moral storm is a very effective frame for understanding that there will be ethical problems to be found when engaging with geoengineering. Yet it does not provide guidance on how these ethical concerns should be evaluated. For example, one may use the perfect moral storm to identify a set of ethical problems which are associated with SSI. This tracks the style of the first wave of ethical concerns which were raised with geoengineering, which tended to solely focus on identifying and understanding ethical problems which exist with geoengineering (Preston, 2016: xi). Yet there is a distinct limitation to this approach, which is that it misses the comparative nature of judgements about geoengineering. It is not enough to make a judgement about the ethics of a

particular response to geoengineering in isolation; such a judgement should involve comparisons with other possible approaches (Preston, 2016: xii). The risk-risk trade off frame is very useful for drawing our attention to the comparative nature of judgements about SSI. By the risk-risk trade-off highlighting this it guides ethical analysis via frames such as the perfect moral storm to help prevent these judgements falling foul of the error of forgetting about the comparative nature of these judgements.

Overall, by combining these frames we create a framework which is simultaneously sensitive to the comparative nature of judgements about SSI, but does not allow the ethical causes of concern with SSI to fall out of sight.

One could accept the point that the risk-risk trade off frame alone invites ethically suspect reasoning, and that we need an ethical framework to mitigate this, yet resist the idea that the perfect moral storm is the appropriate framework. Specifically, one might observe that the thesis has already introduced a normative framework of minimal rights. Perhaps all we need to do is recall this framework when considering the risk-risk trade off and we can see what is at stake in decisions about geoengineering governance, since many of the risks of SSI will endanger these minimal rights. This raises a question about whether we need the additional perfect moral storm frame in order properly to engage with the ethics of geoengineering governance. Indeed, we might be concerned that the addition of theperfect moral storm will confuse our analysis, by introducing two potentially competing normative standards. Yet the mistake lies exactly here, in thinking that the perfect moral storm is a normative standard comparable to the minimal rights framework, and one which serves a similar function. The perfect moral storm is descriptive in nature; it is a description of the challenge of climate change and the challenges we face in trying to address this issue. The great benefit of the perfect moral storm is that its description of these challenges is one which vividly draws attention to ethical problems. Yet it does not provide a normative standard by which to judge whether these problems have been adequately addressed. Therefore, there is no tension between the perfect moral storm framework and the minimal normative standard outlined in chapter 1. The thesis will show how drawing on both frames, together with a commitment to minimal rights, can help us to identify three key ethical challenges in SSI governance, and to begin to think about how to address them

2.5 Conclusion

This chapter has endeavoured to give the reader the tools they need to in order to engage with the following chapters of the thesis. To this end, it has provided the reader with a wider picture of the ethical concerns with SSI, which the problems of moral corruption, moral hazard and intergenerational inclusion sit among. Moreover, it has provided the reader with two frames which can be used to help make sense of what we are doing when trying to address the problems of moral corruption, moral hazard and intergenerational inclusion.

3 How to Reduce the Likelihood of Moral Corruption When Governing SSI

3.1 Introduction

Moral corruption is a problem which our responses to climate change are susceptible to (Gardiner: 2011). A core case of moral corruption is the abuse of power in the belief that such behaviour is morally acceptable. To explore how the risk of moral corruption can be reduced in the governance of SSI, I will first provide an account of moral corruption. This will be followed by the conceptualisation of relevant concepts to address moral corruption, those of transparency, publicity and accountability. Before exploring who these concepts will be used to address moral corruption I will consider the language of geoengineering governance reports to show that these reports have failed to appreciate the relationship between this concepts. The chapter then proceeds by considering whether transparency, publicity or accountability holds any promise for addressing moral corruption. The analysis concludes that accountability does hold the potential to address some symptoms moral corruption, but that accountability does this if the conditions of transparency and publicity being met. The benefit of this analysis being that we now have a possible way of addressing moral corruption in SSI governance.

This chapter also has value insofar as it contributes to understanding how SSI should be governed, independent of our concerns about moral corruption. By exploring the relationship between transparency and accountability in the context of SSI, the chapter shows how poorly this relationship has been understood by SSI governance reports. Such reports tend to highlight the virtues of transparency (Morrow, 2017: 8) but appear to misunderstand its value independent of effective accountability. Notably, accountability is absent from the Oxford Principles, which is one of the most well-regarded set of principles for geoengineering governance (Rayner *et al*., 2013). Consequently, the authors of these reports misunderstand the value of transparency which they place such weight on, and they are less likely to achieve their desired outcomes, for which effective accountability is required.

The relationship between transparency, publicity, and accountability is still neglected and poorly understood, even in contemporary works on the topic. Recently, a report was produced called The Governing Solar Radiation Management Report: Academic Working Group on

Climate Engineering Governance (Chhetri *et al.*, 2018). This report fails to account for this relationship, which is deeply concerning, and it appears that no report has yet been produced which is sensitive to the relationship between transparency and accountability. Moreover, the contemporary literature on the ethical principles of SSI governance also fails to fully appreciate this. Consider the Tollgate Principles as proposed by Gardiner and Fragnière. These principles are designed to explicitly to capture important ethical considerations (Gardiner and Fragnière, 2018: 167). Whilst Gardiner and Fragnière's paper does emphasise the importance of accountability, it fails to appreciate the importance of transparency for accountability to be well-functioning. The value of the current chapter is to highlight how the relationship between transparency, publicity, and accountability should be understood if accountability is going to function well. By doing so, it will draw attention to a systemic error in the literature on SSI governance and show how well-functioning accountability can help address the problem of moral corruption.

3.2 What is moral corruption and why should we care about it in the case of SSI governance?

Stephen Gardiner presents the idea of moral corruption in his book *A Perfect Moral Storm* (2011). Recall that the perfect moral storm provides an account of the challenging situation we are in when responding to climate change. The temptation to succumb to moral corruption is increased by the challenges of the perfect moral storm. Thus, when considering proposals to engage with moral corruption, we should be sensitive to the broader context in which they arise: the perfect moral storm.

Gardiner does not provide a definition of moral corruption, instead he points to the core features which moral corruption possesses.[76] The first of these is self-deception, because the morally corrupt person convinces themselves they are reasoning in a moral way. Self-deception is a complicated idea, in part due to the implication that it requires an individual to believe something they think is false at a particular point in time, or over time (Deweese-Boyd, 2016). Those who hold that self-deception is possible tend to divide on the question as to whether the self-deceit is intentional or not. Those who reject intention as the grounds for self-deception replace intention with other forms of motivation. Both accounts have some intuitive appeal and, importantly, they are both able to tell a story about self-deception.

The second feature of moral corruption is the alignment of self-deception with self-interest (Gardiner, 2011: 307). This is based on a narrow understanding of self-interest rather than an account of an individual's *genuine* self-interest, since that would include an interest in acting morally and avoiding self-deception. Thirdly, the actions committed as a consequence of moral corruption create victims who are unable to effectively resist the behaviour of morally corrupt actors (Gardiner, 2011: 46). In the case of climate change, these victims comprise the global poor, future generations and (arguably) non-human nature.

The fourth aspect of moral corruption is that it shifts the terms of the debate, so that we do not focus on the right types of questions (Gardiner, 2011: 302). Consider, for example, how we speak of animals being harmed by climate change, where the vocabulary is that of species not individual animals. The category of species focuses our attention on keeping certain species in existence or plentiful, rather than considering the harm that animals experience, independent

[76] These core features are not necessary and sufficient conditions of moral corruption but rather they are the features that are associated with core cases of moral corruption.

of the species they happen to belong to. If animals have moral rights then that value is hidden by the debate being conceived of in terms of species (Kapembwa and Wells, 2016). Finally, moral corruption weakens our commitment to moral norms (Gardiner, 2011: 304). This might follow as a consequence of the shift in the focus of the debate, because on that basis such norms are not then taken seriously. Together, these core features of moral corruption present a particularly troubling phenomenon.

There are further concerns with moral corruption, independent of our concern with self-deception.[77] Let us consider an individual's character as well. In the extreme case, someone who is morally corrupt seems to have lost their grasp on what morality is; they might seem condemned to a life devoid of moral behaviour, except perhaps on those occasions where morality and self-interest align. Insofar as thinking and behaving morally is part of a good life, it looks like an individual is unable to lead that life whilst being subject to moral corruption.

As previously explained, moral corruption also has a strong influence on our capacity for moral reasoning. This leads to a problem concerning the quality of public discourse. Moral corruption dilutes the quality of public discourse, to the extent that public discourse is not about what is morally at stake, but about arguments in the name of morality which actually serve the interest of those making the arguments (Gardiner, 2016: 40). An additional reason that we might have for caring about moral corruption is the tendency for moral corruption to lead to undesirable outcomes, such as inflicting harm on the helpless.[78] And indeed, as already noted, we have reasons to be troubled by moral corruption even if it does not lead to such outcomes. The concern of moral corruption is pervasive throughout responses to climate change, given that our temptation for moral corruption is amplified due to the perfect moral storm. The likelihood of moral corruption happening when engaging with a particular response to climate change seems to be contingent on how self-serving that response to climate change can be made. No response to climate change appears to be free from the danger of self-serving interests. Nonetheless we may think that this danger varies between responses to climate change. For example, we may think that reforestation does not present many opportunities to be self-

[77] There are clear consequentialist grounds for being worried about moral corruption, insofar as if one acts on morally corrupt reasoning, it looks as if this will produce harm. However, there are also important non-consequentialist concerns to be found in the self-deception literature, which are relevant to understanding why moral corruption is bad (Deweese-Boyd, 2016).People need not exhibit all of these flaws in the case of moral corruption, although they all seem intuitively plausible. However, taken together, these features of behaviour ground a general argument as to the undesirability of moral corruption.

[78] Given the seriousness of climate change, we want actors to possess a high quality of character alongside a high quality of reasoning, and for this to happen with a high quality of discourse. That is to say, the seriousness of climate change might make us considerably worried about our vulnerability to moral corruption in this context.

serving, therefore the fear of morally corrupt reforestation may be weak. SSI, as explained in chapter 1, definitely provides the opportunity to be self-severing in multiple dangerous and unethical ways, hence we have particular strong reasons to be concerned with moral corruption in the case of SSI.

As has been explained, there are a variety of reasons to be concerned about moral corruption in general. Given the high ethical stakes of SSI, good governance is of particular importance. Even if we have the purest of actors in the governance process, SSI still raises difficult problems. Moral corruption seems to attack the quality of actors – be it by corrupting their consciousness, or by causing them to underestimate their 'courage and self-control' (Deweese-Boyd, 2016). If so, our general concern with moral corruption in the case of governance is that it results in unfit actors forming part of the governing process. Hence it is clear that we have strong causes for concern about moral corruption within SSI governance.

I shall briefly comment on the types of agents who I have in mind and the stage of SSI. The agents who I will focus on being the subject of moral corruption are policymakers. This is for two reasons, firstly that policymakers could be subject to moral corruption given the challenges of justly responding to climate change. Secondly if policy makers are subject to moral corruption then this could be impactful. Now we have a clearer sense of the relevant agents we can consider the stage of SSI. I understand the relevant stage of SSI as that of research and development, although the concern of moral corruption could of course occur at other stages of SSI.[79] I will now explore how the risk of such corruption can be reduced by considering the idea of transparency.

[79] This point may seem underdeveloped, indeed there is much more which could be said about the type of agents who could experience moral corruption and the stage at which is occurs. Yet not much of the analysis is contingent upon these features. Additionally if one desires a more detailed account of who policy makers are, they can insert the account of the formal role holders of power as explained in part 4.5 of this thesis.

3.3 Why consider transparency?

Transparency is a powerful tool. It is argued that it can reduce the risk of war, compensate for the lack of accountability in international organisations and improve regime effectiveness (Lindsedt and Naurin, 2010: 301). This is an impressive list of powers. Additionally, it is an idea which has had broad appeal, with arguments in favour of transparency coming from enlightenment thinkers such as Jeremy Bentham (1931 [1802]), Adam Smith [1776, chapter 2], Jean-Jacque Rousseau ([1772], 1985:72) and Immanuel Kant [1795], as well as Woodrow Wilson, parts of the Christian tradition, and the feminist movements of the 1960s and 70s (Hood, 2006: 6-8).[80] Within the geoengineering governance literature, transparency also appears to receive near universal endorsement (Morrow, 2017: 8). There is empirical evidence which indicates that transparency reduces corruption: bribery, for example (Lindstedt and Naurin, 2010: 317).

This raises two questions: (1) how are moral corruption and corrupt practices related; and (2) why should a solution to the latter be of any relevance to the former? The answer to question one is that corruption takes a variety of different forms, for example bribery and nepotism. We can understand, oral corruption as a distinct subtype of corruption, alongside these more familiar forms. Gardiner explains that, despite the lack of a philosophically rigorous account of corruption itself, we understand corruption as having two core features (Gardiner, 2010: 303). These features are those of distorting or altering for the worst, as well as the corruption being immoral or dishonest and particularly with the use of power (Gardiner, 2010: 304). If we accept these as core features of corruption, the difference between moral corruption and other instances of corruption is where this corruption takes place. Moral corruption can occur when thinking and engaging in discourse about climate change. Moral corruption contains the core features of corruption, insofar as it results in a distortion of thought and dialogue on climate change in such a way that reduces the quality of our moral reasoning. This could, in turn, result in what we recognise as corrupt practices, such as wealthy nations using international climate negotiations as a means to advance their economic self-interest.

Moral corruption is a very useful concept, because it warns about the poor quality public dialogue we can expect to see surrounding decisions about climate change. Yet there are two

[80] The reasons for these agents being interested in varies, for example Smith was interested in transparency due reasons relate to taxation (1776, book 5, chapter 2), whilst Kant was interested transparency in the international arena and the role it could play in helping peace exist between nations (Kant, 1795).

parts of moral corruption which are less clear. Firstly, what is the relationship between poor quality public dialogue and poor decisions? Secondly what is it that really concerns us about moral corruption?

To answer the first question I will propose a hypothesis that will be granted as true during this thesis. We can grant that the quality of decisions is in part an effect of the quality of the dialogue. Therefore, if moral corruption infiltrates that dialogue, then not only will the quality of discussion will be worse, so will the subsequent decision-making. This present provides a simple way to understand the link between dialogue and decisions.

Moral corruption is something we should care about and seek to prevent. Our reasons for this were explained in the previous section: they range from concern about the character of agents who experience self-deception, to the consequences of actions which result from moral corruption. When decision about SSI are made moral corruption has the potential to contribute to poor choices being made.

These outcomes of moral corruption are the focus of this chapter, and I consider principles and mechanisms which could reduce or prevent the negative effects of moral corruption on decision-making about SSI. There are multiple reasons for focusing on this part of the moral corruption problem. One is about our capacity to address the problem: there is a sense in which it is easier to address and regulate the outcomes of moral corruption (such as corrupt behaviour) as opposed to genuinely moral corrupt thoughts or dialogue. The second reason is that some actions which result from moral corruption are harmful and unjust. Therefore, the outcomes of moral corruption need to be addressed.

Given the importance and relative ease of focusing on the outcomes of moral corruption, this is the part of moral corruption which this chapter will seek to address. It is a further interesting question of what measures might affect the moral corruption of thought and dialogue, yet one which is also hugely challenging and one which this chapter will not address.

The answer to the question of why moral corruption are corruption are related is that the two phenomena share core features. As identified by Gardiner, these are: (1) the abuse of power for illegitimate personal gain; and (2) the presence of victims who lack a voice (Gardiner, 2011:

304). Due to these common features, it makes sense to ask whether the ways of engaging with corrupt practices have any impact on moral corruption as well.

Despite the intuitive appeal of transparency, there are also many reasons to be cautious about placing too much trust in the transparency mechanisms. There is the question of how transparency manages to reduce corruption: is it transparency itself which does the work, or is transparency part of a mechanism which produces these results? Lindstedt and Naurin (2010) argue that publicity and accountability are necessary for transparency to reduce corruption, and I will use their framework here. I will show that transparency on its own has little hope of reducing the likelihood of moral corruption and that it is therefore important to consider it in the broader context of publicity and accountability. The reason why there is little hope of transparency reducing moral corruption by itself is because additional mechanisms are required to increase the likelihood of regimes responding to the concerns which effective transparency procedures might prompt. For example, if people want to disapprove of moral corruption (or evidence of it), then suitable mechanisms are required for this disapproval to be expressed. Accountability is one such mechanism, but transparency alone is not.

To consider transparency, publicity and accountability it is useful to have an account of the relationship between these three concepts. There is a debate in the contemporary literature about how transparency and accountability ought to be conceptualised and the relationship that these conceptions stand in to one another (Naurin, 2006; Fox, 2007; Hood, 2010; Lindstedt and Naurin, 2010; Meijer, 2014; Mabillard and Zumofen, 2017). Unfortunately such literature neglects the role of publicity,[81] therefore I shall account for publicity after considering the relationship between transparency and accountability. Christopher Hood identifies a helpful approach to thinking about the relationship between transparency and accountability concepts (Hood, 2010). He argues that there are three different understandings of the relationship between transparency and accountability, which he names; Siamese twins, matching parts, and awkward couple (Hood, 2010). The Siamese twins account argues that transparency and accountability are interchangeable terms for the same idea (Hood, 2010: 990). The matching parts account understands them as distinct concepts which are complementary to each other and produce good governance (Hood, 2010: 992). The Awkward couple account also acknowledges that they are separate concepts but fears that they concepts have the potential to produce low quality governance, for example when there is transparency agents will be

[81] The exception being Naurin (2006) and Lindstedt and Naurin, (2010), both of which engage with publicity.

concerned about avoiding blame as oppose to taking the risks to do what is good. There is not a general conceptual or empirical relationship between transparency and accountability in all case, instead different relationships will exist in different instances of transparency and accountability (Hale, 2008; Hood, 2010; Meijer, 2014; Mabillard and Zumofen, 2017).

I shall consider why and how Hoods matching parts understanding of the relationship between transparency and accountability. The reason for this is that the matching parts conception of their relationship is sensitive to the idea that transparency and accountability are distinct concepts, this is not true of the Siamese twins account which argues that they are the same. The idea of them being distinct concepts has initial plausibility and bears out in the analysis in this chapter. Moreover the matching parts account of their relationship has the benefit of being implicit in much literature and practice on good governance. There is a justification which is that this account seems the promising, if it can be made to work in the case of SSI governance, it should contribute to the good SSI governance in general and is the account which I think may be the most use for addressing moral corruption. This is the reasons for choosing the matching part over the awkward couple account which brings the opposite promise of poor governance.

Broadly speaking transparency refers to the accessibility of information, whilst accountability refers to the possibility of sanctioning agents. Based on these broad understands of the term we can make sense of the idea that there is a missing concept if we are to achieve the matching parts relationship between transparency and accountability in this case. That concept is publicity which broadly refers to the idea of transparent information standing a reasonable chance of reaching the relevant agents (Naurin, 2006; Lindstedt and Naurin 2010).

We can now move on from the broad conceptions and ask the specific question, how should the terms transparency, publicity, and accountability be understood? There is vast contemporary literature on these question. When considering transparency there is much that one could disagree about (Heald, 2006; Etzoni, 2010; Mabillard and Zumofen, 2016) for example does transparency require that is possibility of being able to access information, such as via freedom of information requests (Michener, 2011). Or does it require the active publication of such information (Mabillard and Zumofen, 2016: 2). What is it that needs to be transparent, an event or the process which produces the event, and the for who does transparency need to be achieved (Heald, 2006). I shall adopt an understanding of transparency which is helpful for the matching parts conception of transparencies relationship with

accountability. I understand transparency in what Mabillard and Zumofen (2016), term active transparency. Whereby the relevant institution voluntarily publishes relevant information about its decisions and the processes by which such decisions are made.[82] This conception of transparency is useful for the matching parts account of the relationship between transparency and accountability due to it insuring that the relevant information is available to hold the relevant agents to account.[83]

Given that transparency is understood as the publication of relevant information, publicity builds on this to require that this information has a reasonable chance to reach the relevant agent and that the information can be easily understood by that agent (Naurin, 2006; Lindstedt and Naurin; 2010).[84] To be clear, this means that publicity has two components, one of which is that there is a reasonable chance of the information being received by the relevant public, secondly that it is reasonable to expect them to understand this information. Publicity is only achieved when both of these conditions are met. Transparency is a necessary condition for the achievement of publicity, given that any instance of publicity will require that the relevant information is published. The distinction between transparency and publicity is useful due to the fact that transparency alone does not appear to ensure that the information is accessible.

The distinction between transparency and publicity is illustrated well by the following case from *The Hitchhiker's Guide to the Galaxy*. In the opening chapter of Douglas Adams' science-fiction novel, an official from the local council comes to the house of Arthur Dent in order to knock it down. When Arthur protests about the demolition he is told that the plans have been publicly available to view and contest for many months. Arthur's terse reply provides us with a humorous insight into transparency and publicity.

> "On display? I eventually had to go down to the cellar to find them."
>
> "That's the display department."
>
> "With a flashlight."
>
> "Ah, well, the lights had probably gone."
>
> "So had the stairs."

[82] There is of course scope for variation in how sincerely an institution does this, for example we might think of an institution as translucent if it publishes some relevant information about why a decision was made but withholds vital information.

[83] This conception of transparency is also adopted in Buchanan and Keohane (2006: 427) to describe what transparency is for global governance institutions.

[84] I understand the comparative lack of literature on publicity as in part a product of the how some definitions of transparency seem to include a publicity condition (Cotterrell 2000; Grimmelikhuijsen and Welch 2012).

> "But look, you found the notice, didn't you?"
>
> "Yes," said Arthur, "yes I did. It was on display in the bottom of a locked filing cabinet stuck in a disused lavatory with a sign on the door saying 'Beware of the Leopard." (Adams, 1995: 19-20).

Is there transparency in this case? The plans to knock down Arthur's house are in the planning department. If Arthur was to go into the planning office and overcome the respective obstacles inside the department, he would be able to find a piece of paper which clearly states that his house will be demolished on a particular date. The plan is transparent, insofar as the plan can be accessed by Arthur. Yet we are clearly troubled by Arthur's case, because though there is transparency there is an absence of publicity; the information is not communicated in a meaningful sense. Information has to be communicated so that the relevant public have a reasonable chance of accessing and understanding it. In this case, it is clear that there was no reasonable attempt to communicate this information clearly to Arthur. Therefore, whilst transparency is present, publicity is absent.[85] This case is troubling and humorous and absurd, and this is all due to the absence of publicity.[86]

This account of transparency may appear to be odd, and the oddness of it was highlighted by an objection from one of my examiners. They asked whether on this understanding of transparency their diary would be considered transparent. Their diary is also hidden in the safety of their home, but if someone were to overcome the relevant obstacles (perhaps comparable to the obstacles Arthur Dent faced), then they could access the diary. It would clearly put pressure on this account of transparency if it understood the diary as transparent. Here I will remind the reader of one of the core features of transparency which is that information is published, *in the sense that it is put in a place where it is legally permissible for the general public to access it.* The reason why the diary does not meet the relevant conditions of transparency is that it is not published. The plans to knock down Arthur Dent's house are published, it is just that this published information is hidden from public sight. Part of the value of this account of transparency is that it helps us understand what is important*;* it

[85] One might object and argue that a broader understanding of transparency would include the publicity condition. This broader view is held by some, such as Cotterrell (2000); Grimmelikhuijsen and Welch (2012). Yet those who hold this broader view would have to provide an account of what happens in the hitchhikers guide case, where by information is obtained despite the lack of transparency.

[86] This thought can be also be seen in the work of Onora O'Neill on transparency, who argues forcefully that transparency alone does not ensure good communication (O'Neill, 2006: 82).

is not enough for information to be published, we also want something akin to the publicity condition to be met.

The last term to define for this analysis is accountability. As is the case with transparency, there are a variety of ways in which accountability can be understood. Koppell (2005) argues that this is a product of many approaches one could adopt when studying accountability (see also Mabillard and Zumofen, 2016).[87] Accountability is broadly understood as denoting an agent or institution who is answerable for their affairs (Hood, 2010: 989).[88] Mabillard and Zumofen (2016) observe that Hood's (2010) understanding of accountability is consistent with a three stage process of power transition outlined by Boven (2007, 454). The three stages are:

> 1: The actor is obliged to inform the forum about his or her conduct (information phase)
>
> 2. The forum can question the adequacy of the information or the legitimacy of the conduct (debating phase)
>
> 3: The forum may pass judgment on the conduct of the actor (consequences phase).' (Boven, 2007:454)

One of the main divides in the accountability literature is the third stage, specifically whether the forum requires formal sanctioning power. Those who adopt a soft account of accountability do not think that formal sanctioning power is required; instead they think that accountability occurs when an agent is asked to answer for their actions, even if the agents asking lack formal sanctioning power, such as the media (Fox, 2007: 668). On the other hand, there are those who think answerability is not enough and that formal sanctioning power is also required; this is known as a hard conception of accountability (Fox, 2007: 668). Both of these forms of accountability have a proper place; it is appropriate that the media does not have formal sanctioning power, but nonetheless holds relevant agents to account by scrutinising and publishing their actions. Yet hard accountability also has an appropriate place. A clearly desirable circumstance for hard accountability is when such sanctioning power can legitimately be used to incentivise agents to 'behave themselves'. I am going to accept the premise that it

[87] Regarding whether accountability is a virtue or a mechanism, I shall understand it as a mechanism, which is not to deny that it can also be a virtue.

[88] Accountability can be understood both as a principle and a mechanism. The mechanisms of accountability would refer to how an agent is held answerable for a state of affairs which they produce, whilst the principle of accountability refers to the ideal of accountability. The mechanism is a way in which this principle can be achieved. Therefore, we could have different mechanisms for achieving the same principle of accountability, or different mechanisms for achieving different understandings of the principle of accountability.

is legitimate for SSI governance to sanction agents who partake in SSI governance (this point is explained in more detail in section 3.7 of this chapter), therefore I will adopt a hard understanding of accountability

3.4 The argument about transparency in the context of SSI governance

I noted earlier that the relationship between transparency, publicity, and accountability is often poorly understood in geoengineering governance reports, and that this limits the effectiveness of the recommendations that these reports produce, at least in regard to transparency. As stated, the relationship with publicity and accountability is vital for understanding the force of transparency. I will now try to present the relevance of this argument by showing how it seems that transparency is often conceived of without any consideration of accountability in SSI governance reports. In the Royal Society report we read:

> The Royal Society in collaboration with international science partners should develop a code of practice for geoengineering research and provide recommendations to the international scientific community for a voluntary research governance framework. This should provide guidance and transparency for geoengineering research, and apply to researchers working in the public, private and commercial sectors. (Shepherd *et al*., 2009:p. xii)

Additionally, a report by the US Bipartisan Policy Center (2011) states:

> Principle 4: Importance of Transparency. Federal officials and federally funded scientists must keep the public informed about CDR and SRM research, especially when that research itself could be harmful or the method could be implemented at scale without substantial effort. Research plans and research results, both positive and negative, should be made public. The federal government should develop transparency protocols for all potentially risky forms of climate remediation research. Those protocols should be appropriate for the magnitude and extent of potential impacts for the specific experiment under consideration – that is, protocols should be based not only on the risks posed by related research, but also on the risks that would be posed by deployment. (Bipartisan Policy Center, 2011: 14)

Both of these quotes highlight the importance of transparency, and both fail to mention its relationship with accountability.[89] This is true all the way through both of these reports, in that accountability is not mentioned at all in the Royal Society governance report, and only twice in the Bipartisan Policy Center report (although one of these uses is in naming a reference, so to be strictly accurate the idea of accountability is only mentioned once in the main text).
The same is also true of the following reports: The Congressional Research Service Report (2010); the Kiel Earth Institute's 'Large-Scale Intentional Interventions into the Climate System? Assessing the Climate Engineering Debate' (Rickels, *et al*: 2011); 'Solar Radiation Management: The Governance of Research' by the Solar Radiation Management Governance Initiative (SRMGI) (2010); 'Geoengineering for Decision-Makers: Science and Technology (2011); the Ecologic Institute Report (2014); the National Academy of Science (NAS) Report (2015); the EUTRACE Report (2015); and the report of the Academic Working Group on Climate Engineering Governance 'Governing Solar Radiation Management' (Chhetri *et al.*, 2018).[90]

We have yet to see a report which seems to have a clear grasp of the relationship between transparency and accountability! Despite this, the idea of transparency has received a wide variety of endorsements, as is shown by following quotes. These quotes also show that there is some confusion about how transparency ought to be operationalised. I understand this to be partly due to confusion about how transparency could be put into practice in a manner similar to other concepts such as accountability.

> There is consensus on the need for transparency, but there still is much disagreement over the scope and design of a disclosure mechanism. (Turkaly *et al.*, 2017: 1)

> However, there was broad agreement that a well-designed transparency mechanism would provide a common, central source of current information and is necessary, if not sufficient, to ameliorate polarization and inform the climate policy conversation. (Turkaly *et al*., 2017: 4)

[89] Publicity is also absent from these quotes, yet that seems less surprising and noteworthy given its specific and contested technical nature
[90] These reports tend to be produced by government agencies, or non-governmental organisations.

> If research does proceed, transparency and openness are critical. The reports agree that CE [climate engineering] research, and especially research in solar geoengineering, must be transparent and open, rather than proprietary or secret. The reports do not always specify or agree about exactly what transparency requires. (Morrow, 2017: 8)

> 4. How should transparency be operationalized? While the need for transparency is among the most commonly cited principles for CE research, the reports say relatively little about how to operationalize this principle. The primary literature contains some more detailed suggestions. (Morrow, 2017: 12)

It is clear, thanks to Morrow (2017:8) that transparency is endorsed by nearly every geoengineering governance report. And, also, that publicity and accountability is neglected by nearly all of these reports. I will show that this is seriously problematic, at least in the context of transparency as a tool to engage with moral corruption. Moreover, this is also likely to be problematic if transparency is to be a meaningful tool to engage with any problem of governance. If I am correct, I will show that many of the geoengineering governance reports have failed to recommend what they would desire because they do not understand the relationship between transparency, publicity, and accountability.

3.5 What does transparency have to offer in the case of moral corruption?

What can transparency do to engage with moral corruption? An analysis of this question will show that transparency is not the whole story in regard to moral corruption and that accountability is going to be vital to reducing the risk of it. However, transparency does have significant instrumental value as a means of holding actors to account for moral corruption, as I shall explain.

When considering transparency we need to consider not just the information which is available, but also who will be able to access this publicised information. Transparency in relation to an SSI global governance regime can clearly be achieved for at least one set of people throughout the world, namely, those with internet access to global governance regime websites or internet access to those who report on global governance activities, such as non-governmental organisations (NGOs) and research councils. As I have explained, transparency is based on the availability of information, and that information can be made freely available on the internet.[91]

However, would this lessen moral corruption policymakers? The answer depends on at least two things. Firstly, there is an empirical question in terms of whether knowing that people can access this information would lead to a change in the type of reasoning which actors in SSI governance engage in. There are reasons for thinking that this might be the case and there is a whole literature about how people's behaviour changes when it is subject to observation (a simple example is that of how people's behaviour changes when washing their own hands if they think they are being observed (Sunstein and Thaler, 2008)), but is this observation enough to ensure a change in behaviour in cases of moral corruption?

An important element in regard to moral corruption is the move towards an agent's narrowly construed self-interest, and how they maintain or even strengthen their sense of moral integrity.[92] Therefore, if the morally corrupt agent fails to recognise there is anything wrong with their morally corrupt state, then the prospect of the public becoming aware of it will not motivate a change in behaviour. Also, even if they do recognise something wrong with it, it is nonetheless unclear how they will respond. Given the lack of sanctions and the strength of self-interest, it seems foolhardy to expect there to be a substantial decrease in moral corruption.

[91] Examples of the kind of information to which I am referring will be provided in the next section, in which I apply transparency concepts to some of the domains of SSI governance.
[92] The sense of moral integrity is achieved because the actor thinks their behaviour is moral.

The second, related, reason for doubting the impact of transparency concerns who the information becomes transparent to. If information only becomes transparent to those actors who would benefit from the relevant decision, then it is questionable whether that would lead to any change in behaviour by the regime. However, if the behaviour of a regime becomes transparent to those who are vulnerable to the actions of that regime, then the regime may be more likely to change its behaviour so that it does not appear to be harming those who are now aware of its actions. The extent of this change in behaviour may be contingent upon the powers of the vulnerable. This is also conjecture, based on the assumption that the regime in question has some reason to care about how it is perceived by those groups who are affected by its actions, and that the regime in question is responsive to certain actors.

I will now consider whether transparency can be achieved for those who are truly vulnerable to the effects of moral corruption in this situation. Whilst it is an open question as to whether the governance of SSI can be made genuinely transparent to the global poor, clearly this cannot be done for future generations.[93] Therefore it is unclear what transparency can do for vulnerable people if this information cannot be made clear to them. Similarly, if transparency is understood as the provision of information, then it seems to be impossible to provide it to future generations at this point in time.[94] This also brings us back to the question of why transparency *alone* would lead to a change in behaviour. Unless actors care about how they will be perceived over time, it seems unlikely that this will lead to a change in behaviour.[95]

To summarise. For transparency alone to be effective, it would need to be achieved for vulnerable groups. However, transparency cannot be achieved for all those groups who are vulnerable to the effects of moral corruption in an SSI governance regime. Therefore, it appears that transparency on its own will not reduce the risk of moral corruption.

[93] Given the barriers which the global poor face in terms of poverty and quality of education, it is clearly a challenge to make information about a global governance regime transparent to them. Moreover, it is doubtful whether it would be appropriate to try to make SSI governance transparent to people who face such barriers without at the same time providing action to alleviate them.

[94] Things could be made transparent for future generations upon their temporal arrival – the preparation of information which will become available to people at a certain point in the future – but this removes a lot of the transformative potential of transparency in the here-and-now.

[95] This is based on the premise that relevant actors do not currently care enough for the *interests* of future generations, although they may care about their *relationship* with future generations. For a possible explanation of this, see McKinnon on moral shame (McKinnon, 2012: 134–135).

3.6 Transparency and its friends: considering publicity and accountability in SSI governance

I will now consider two concepts which are related to transparency: namely, publicity and accountability (Naurin, 2006). If transparency is concerned with the *provision of information*, then publicity is restricted to the *scope of transparency* and whether people understand the information that is available to them.

As was established in the previous section, it seems doubtful that transparency alone would lead to a change in an actor's behaviour or reasoning, unless that actor also cared about how they are perceived. Whilst the condition of publicity means that those who receive information as a result of transparency ought to be able to understand that information, it does not obviously provide a reason as to why an actor may care about how they are perceived. Moreover, neither does it seem to provide any additional reasons for reducing the risk of moral corruption happening. Given this, I am forced to draw the same conclusion as I did from my analysis of transparency: publicity alone is not a useful tool for engaging with moral corruption.

When thinking of the role of publicity in this case it is worth drawing to attention to what is required for the publicity to be met. Information about SSI has to have a reasonable chance of reaching the relevant public and it has to be easily understandable to them. Both of these conditions may be demanding. The demandingness of the first conditions is in part contingent upon how the relevant public is understood, if it is broadly understood this could require that the majority of people in the world need a fair chance of receiving this information. Again this raises general questions of whether enough people across the world have the appropriate means to access this type of information. Moreover this relevant public has to find this information easy to understand. This raises a huge question about what information is meant to meet the publicity condition. The following consideration can this information be made easily understandable. This question presumably turns on considerations such as how well this information needs to be understood for the publicity condition to be meet.[96]

[96] The feature of demandingness would not be nearly as strong in the case of achieving active transparency. As explained, active transparency does not require that the information has a fair chance of reaching the public or that is easy for them to understand., Merely publishing the relevant information on the internet may be enough to satisfy the demands of active transparency.

Let us turn then to accountability. In the context of SSI governance and moral corruption, accountability is concerned with the ability of people to impose sanctions upon institutions. Sanctions are a tool used to influence behaviour, in the knowledge that they can be used to punish a certain type of behaviour. The scope of transparency is not necessarily the same as the scope of accountability, and as Christopher Hood explains, you can have accountability in the absence of transparency (Hood, 2010: 992–993). He illustrates this by pointing to the example of how civil servants can be held to account by ministers, and how intelligence services can be held to account by legislatures, yet transparency can be absent or limited in both these cases (Hood, 2010: 992-993).

We also know that accountability can lead to an institution changing its reasoning and behaviour. Consequently, we can conclude that the type of accountability which is able to hold an institution to account for how its behaviour affects those who are most vulnerable to it is the type of accountability that could lead to a positive change in reasoning and behaviour from that institution, not only in regard to present generations, but also in regard to future generations and non-human nature.

Accountability is clearly very important here. However, we need to pause and consider how it is that accountability has such force. This is linked to publicity and transparency. The idea of publicity is concerned with understanding the information that an institution produces. If one wishes to hold an institution to account effectively, then it would appear that one first needs an understanding of how that institution has behaved and why it has behaved in a particular manner.[97] However, this information would not be available without there also being transparency. Therefore, in terms of the account of transparency, publicity and accountability provided here, transparency and publicity are of huge instrumental value to accountability, and accountability is the way in which the likelihood of moral corruption is reduced. This also shows that faith in transparency *per se* is misplaced, and that the power of transparency should rather be understood in relation to how it enables processes like accountability to function effectively.

The question to ask here, then, is: how does holding an institution to account on behalf of vulnerable groups reduce the risk of moral corruption? Whilst it is clear that accountability can

[97] The condition of publicity – at least, the minimally acceptable scope of publicity – has to be met for those who are meant to hold the governance institutions to account. However, this should not necessarily be considered the optimal level of publicity. There are probably strong reasons to say that the scope of publicity should be greater than it is, but the outline here relates to the minimum rather than the optimal level of publicity.

affect the behaviour of those who are morally corrupt, it is less clear how it can engage with those experiencing corruption. That being the case, it appears that accountability engages merely with a symptom of corruption and not the problem itself. If we consider self-deception, of which moral corruption is an example, we can identify two general concerns:

> Self-deception violates an epistemic duty to properly ground self-ascriptions (Fernández 2013) or violates a general duty to form beliefs that conform to the available evidence'. (Deweese-Boyd, 2016)

When an actor can be held to account via a process of transparency, they will presumably be incentivised to present themselves in such a way that sanctions are not implemented against them. We can assume that one of the standards against which their statements will be measured is that of consistency with the publicly available evidence. If this is the case, it is likely that some actors who might have been vulnerable to moral corruption would not now succumb to it, given the presence of these conditions of transparency and accountability.[98]

[98] Of course, there are relevant questions which have not been answered here, most notably: what type of sanctions are required to change the behaviour of an SSI governance institution? Whilst this question will not be answered here, it is clear that an answer to it is required if this logic is going to be applied to SSI governance.

3.7 An illustration of accountability in SSI research and development

So far, this chapter has argued that accountability is a way to address the issue of moral corruption – given certain conditions being met – such as transparency and publicity. Up to now this has been quite an abstract argument (although it has drawn on some 'real-world' examples such as that presented in the *The Hitchhiker's Guide to the Galaxy*), which is helpful for clarifying the theoretical distinctions between transparency, publicity and accountability. Moving away from the abstract to a concrete example, I will now try to illustrate how accountability (specifically) could be designed for SSI governance by considering how an international accountability mechanism currently functions. This is the mechanism adopted by the Organisation for the Prohibition of Chemical Weapons (OCPW). By looking at how this mechanism functions I will show how we could make sense of an accountability mechanism for SSI governance.

Prior to considering how the accountability mechanism in the prohibition of chemical weapons functions, it is worth considering who should be held accountable for SSI research and development. In other words, which type of agents do we wish to hold to account? Is it the scientists who perform the research, the person or institution which funds the research, the institution where the research takes place, policymakers in a state, or the state which allows it take place? (It should be understood that this is not an exhaustive list of those whom we may wish to hold to account. Importantly, it is not even at all clear which agents ought to be held to account.)

A common intuition for accountability is that the person who performs an action can be held to account for the performance and consequences of that action.[99] I take this to be what Jamieson means when he refers to a notion of responsibility which is a product of common sense morality.[100]

[99] It is worth noting that there is a difference between whom we may wish to hold accountable and who should be held morally responsible. This is expressed by Duff as well as Hart, who explains that the individual whom we wish to hold legally responsible is not necessarily the one who is morally responsible or blameworthy (Hart, 2008: 217; Duff, 2007: 20). The reason that Hart offers for this is that the substance of legal and moral rules is different (Hart, 2008: 225).

[100] Of course, responsibility and accountability are not synonymous concepts, but they are clearly strongly related. When someone is held legally responsible for something, it appears that a condition has been met to hold that person legally accountable as well.

I shall now consider a simple case to show why we have reasons to be concerned about this account of common sense morality in the case of SSI. We need to ask whether it is appropriate for individual scientists to be held to account to a global SSI governance body. One might think that scientists are not the appropriate type of agent to be held to account by a global governance institution.[101] Such a view would require justification. Nonetheless, it is not uncommon for the international arena to be selective about who it is willing to hold to account. Consider for example international agreements by states where only states are liable under them.[102] Therefore, a compelling reason that could be offered is that scientists do not have the appropriate power to compensate for the wrongness of their action. This type of reasoning can be seen in Joel Feinberg's discussion on collective responsibility (Feinberg, 1970: 228–229).[103] The resources available to scientists are much more limited than those available to other types of actors such as states. Therefore, a state is more likely to have the appropriate power to provide compensation.

There are various forms of responsibility which are designed to hold an agent responsible for the actions of others. One such form is vicarious responsibility, as defined in this extract by Feinberg:

> There is vicarious liability, on the other hand, when the contributory fault, or some element of it, is properly ascribed to one party (or group of parties), but liability is ascribed to a different party (or parties). In such cases we say that the latter party is responsible for the harmful consequences of a faulty action or omission of the former party. The person who did or caused the harm is not the one who is called upon to answer for it. (Feinberg, 1970: 226)

As Feinberg's quote shows, fault could be given to an individual agent such as a scientist, but liability could be assigned to a different agent, such as the employer of the scientist. This is useful for the person who would be troubled by the idea of holding the scientist to account. Nonetheless, I take this to be problematic, given that we are considering a global governance institution. I think that we have reasons to consider the state as the relevant actor to hold to account, such as the state having the appropriate power to remedy some of the harms from

[101] In this case scientists are illustrative of any agent who is not the state, you may hold view indicated earlier in this chapter that policymakers are the relevant agent instead, and the same question would arise whether they are the appropriate agent to hold responsible by an SSI global governance institution.
[102] A reason for this is that states are more likely to have appropriate power to correct the wrongness of their action.
[103] These reasons could of course be related: a reason for holding states to account may well be that they are the type of agents who tend to have the appropriate power to compensate for harms.

illegal SSI research and development. Yet this is not the end of the story, it is at least expected of a state to create a domestic legal framework which is consistent with international law. Therefore, a state may take action domestically against an agent who risks the state being held vicariously responsible by using domestic legal instruments. This is the first lesson from international accountability that I will draw attention to: it is common practice to hold states vicariously responsible for actions which are performed within that state, even if they are not authorised by that state.

The idea of vicarious responsibility is clearly helpful for having a clearer idea of how an accountability mechanism could function for SSI governance. But it still leaves an important question unanswered. Namely, *how* should the relevant agent be held to account? Returning to my starting point in this section, I shall now consider how the Organisation for the Prohibition of Chemical Weapons holds actors to account. This organisation is an institution which is a product of the Chemical Weapons Convention which seeks to prevent the use of chemistry in warfare. The reason for considering the accountability in the organisation for the prohibition of chemical weapons is that the mechanisms that the organisation has are ones which do exist in the international arena. In this sense, we could understand these mechanisms as showing what type of accountability mechanisms could be achievable in the international setting. To be clear, this comparison is not based on the premise that chemical weapons and SSI are similar. Far from it. Chemical weapons are weapons (a tautology, but one which is worth stating), and SSI is not a necessarily a weapon.[104]

There appear to be transparency and publicity mechanisms built into the Organisation for the Prohibition of Chemical Weapons. The organisation's Technical Secretariat covers the inspection and verification of state parties conforming to the OPCW's Convention. This requires at least part of the Technical Secretariat to have access to the relevant information from state parties in order for the inspectors to confirm that the standards of the Convention are being conformed to. Transparency requires that this information can be made sense of once the inspectors access that information. Publicity is achieved for the Technical Secretariat if the relevant information is easily accessible for them. Of course, the set of actors to which transparency and publicity apply is very limited. Neither aspect is achieved for the general public in this case.

[104] Although it could clearly be used in an aggressive or military manner.

The Conference for State Parties of the OPCW seems to exercise the accountability mechanisms that the organisation has in place. The Conference has the power of:

> restriction or suspension of a State Party's rights and privileges under the Convention; the recommendation to States Parties of collective punitive measures in conformity with international law and in cases of serious damage to the object and purpose of the Convention; and, in cases of particular gravity, bringing the issue to the United Nations General Assembly and the Security Council. (Organisation for the Prohibition of Chemical Weapons, 2019)

These are clearly accountability mechanisms. Recall that accountability exists when sanctions can be applied to the agent who is meant to be held to account. Depriving states of rights and privileges is a form of sanctions. Additionally, it appears that Conference is in a position to recommend collective sanctions to state partners and to the United Nations General Assembly and Security Council. It is perhaps disappointing to note how weak these sanctions appear to be. The ability to deprive a state of rights and privileges sounds strong, but the rights and privileges which can be restricted are only those which are conveyed by the institution. Therefore, the force of these sanctions is at least in part contingent upon how the sanctioned state values the rights and privileges that are conveyed by the OPCW. Whilst these benefits may be desirable, it seems doubtful that depriving a state of them comes even moderately close to being appropriate as a proportionate sanction to hold a state to account. Nonetheless, despite the arguably weak nature of this accountability mechanism, it does exist, and it does show that this type of mechanism can be agreed upon by the vast majority of states.

The OPCW does have the potential for additional sanctions. These are the ability to recommend 'collective punitive measures' (Organisation for the Prohibition of Chemical Weapons, 2017: 1), and the ability to bring issues to the attention of the United Nations General Assembly and the Security Council. The General Assembly and Security Council are both institutions which have significant sanctioning power. Therefore, the OPCW can have a causal role in other sanctions occurring. We can imagine a comparable mechanism in the case of SSI governance. That is, such an institution could bring violations of SSI treaties to the attention of the UN General Assembly and Security Council, and by so doing, expose the offending agent to the possibility of sanctions imposed by the General Assembly or the Security Council.

This raises the question about what type of accountability mechanism would be appropriate for an SSI global governance institution, and this will be the focus of the discussion which follows. It is disappointing – although perhaps unsurprising – to note that SSI governance reports do not propose how accountability mechanisms could function in the case of SSI governance.[105] If one were to adopt the type of accountability mechanism presented by the OPCW, then an SSI global governance institution would be able to impose sanctions upon its members by restricting their access to, or denying all the benefits which are associated with that institution.

This therefore raises the question as to what type of rights and privileges could be granted to members of an SSI governance institution. A full account of these rights and privileges cannot be provided here, due to the rights and privileges of being part of an SSI governance institution being determined in part determined when an institutions is created. Yet we can predict a minimal account of these privileges. An SSI governance institution could deprive a state of the right to partake in decisions which are made in that institution. We can conceive of much stronger powers that the institution could have, just imagine if membership of other beneficial international institutions (such as membership of the United Nations) were contingent upon being part of an SSI governance regime. Other possible rights and privileges include the ability to make recommendations to the relevant UN agencies to inflict economic sanctions upon that state, or even for there to be a military response. This is illustrative of the type of powers which they can have. I am not advocating that UN membership should be contingent upon SSI governance, but such a governance institution would certainly have much more sanctioning power if it were able to deny members access to the SSI governance institution, and by doing so access to the UN as well. Although the effectiveness of such sanctions would be contingent upon the value of the rights and privileges that an SSI governance institution could provide and deprive a state of.

To conclude this section, there are accountability mechanisms which currently exist in the international system that could be applied to the case of SSI governance. The core lesson from the example of the Organisation for the Prohibition of Chemical Weapons is that the sanctioning power available is that of restricting access to the benefits that membership of the

[105] To be more specific, this is clearly true of the following reports: The Congressional Research Service Report (2010); the Kiel Earth Institute's 'Large-Scale Intentional Interventions into the Climate System? Assessing the Climate Engineering Debate (Rickels *et al*, 2011); 'Solar Radiation Management: The Governance of Research' by the Solar Radiation Management Governance Initiative (SRMGI) (2010); 'Geoengineering for Decision-Makers: Science and Technology (2011); The Ecologic Institute Report (Bodle, *et al;*2014); the NAS report (2015); The EUTRACE Report (2015); and the report of the Academic Working Group on Climate Engineering Governance 'Governing Solar Radiation Management' (Chhetri *et al.*, 2018).

organisation provides. As explained, the force of this sanction would seem to be contingent on the benefits of participating in an SSI governance institution, and it is unclear what those benefits would be or how appealing they would be.

In addition to the question of how an accountability mechanism might function for SSI governance, there is the question of who the relevant agent is. As I have explained, we have reasons for not wanting scientists to be subject to sanctions from an SSI governance body. I have also explained that it is common practice to hold states vicariously responsible for the acts of agents within that state. This seems even more appropriate when we consider that the sanctioning power of international regimes is to deprive members of the benefits provided by these regimes. Those benefits are provided to member states rather than individual scientists. Therefore, restricting the benefits which a regime provides to a state is not a meaningful way to sanction an individual scientist.

3.8 Conclusion

I propose that accountability should be a principle of SSI governance in light of the problem of moral corruption. Yet this comes with a caveat, which is that efforts must be taken to ensure that the mechanism of accountability is well functioning. This requires an understanding of how accountability relates to other concepts such as transparency and publicity. This chapter has provided an account of this relationship in the context of SSI research. Additionally, to ensure that accountability is well functioning it is necessary to understand how accountability can be achieved in the various relevant domains of SSI governance.

David Morrow observes from the reports on geoengineering governance that transparency is consistently recommended (Morrow, 2017: 8). However, he also observes that there is no consensus on how the term should be operationalised. This chapter has endorsed the consensus that transparency is important, due to its relationship with accountability. Therefore, when operationalising transparency, we should do so with an eye to accountability, because after all, accountability is one of the tools we want to apply in the case of moral corruption.

However, accountability appears to be neglected as a principle in proposals for geoengineering governance. For example, it is not mentioned by name in Morrow's summary of geoengineering reports. Moreover, it is notably absent from the most significant set of proposals for geoengineering governance, namely, the Oxford Principles (Rayner *et al*., 2013). Given the near consensus on the value of transparency and the lack of attention paid to accountability, this raises a question as to whether these reports really understand the force of transparency independent of ideas such as accountability, which they seem to neglect. By highlighting the relationship between transparency and accountability, hopefully the conversation on transparency in geoengineering governance can progress with a better understanding not only of why we may value transparency, but also why we may want to give additional thought to accountability.

Finally, if moral corruption is to be taken seriously, then accountability should not be seen as the solution on its own. Nonetheless, it is still more useful that transparency. At best, accountability is a way to engage with the problem. This is in part due to a particular feature of moral corruption, which is that it concerns an aspect of our reasoning process. Whilst tools like accountability may affect reasoning, they are probably better at engaging with the outcomes of such reasoning and negating those outcomes, rather than the reasoning itself.

Therefore, given the difficulties of addressing moral corruption, accountability should be subject to scrutiny before being used as part of any arsenal to tackle moral corruption in SSI governance, and scholars should continue to consider the alternative ways in which moral corruption can be reduced.

4 Understanding the Moral Hazard Complaint: Addressing Incompleteness, Ambiguity and Vagueness

4.1 Introduction

The purpose of this chapter is to provide a clear account of what the moral hazard concern entails when it is applied to SSI, so that the proposals to address the hazard can then be thoroughly analysed in the two chapters which follow. In the broadest terms, the moral hazard complaint is the fear that there will be a reduction in mitigation efforts due to SSI. I will begin by presenting the concept of the moral hazard complaint in relation to SSI, for which I draw upon the work of Ben Hale (2012), Adam Corner and Nick Pidgeon (2014), and David Morrow (2014) in applying the moral hazard concern to SSI.[106] The work of these authors draws attention to the three main components of a moral hazard complaint. Hale's work is helpful for understanding *what* the moral hazard complaint is, Adam Corner and Nick Pidgeon's work is useful insofar as it considers the question of *who* experiences the hazard, and Morrow's relevant contribution is that of thinking about *how* the hazard may happen. When exploring the insights from these authors, it will become clear that two more considerations are necessary in order to provide a complete account of the hazard; namely, the *specific stage* of the SSI process at which it occurs, and *why* the hazard would be bad. Consequently, this chapter will argue that a complete account of the hazard needs to include five variables: what the hazard is, the stage of SSI at which it occurs, who the relevant agent is, the mechanisms by which it occurs, and the impact of the hazard.

These five variables have not been fully explored within the same pieces of work, yet an answer to each of these questions is useful if one is going to make sense of the moral hazard complaint and consider ways to address the problem. Consequently, the contributions by Hale, Corner and Pidgeon, and Morrow all have significant gaps if we wish to use their accounts of the hazard for this project. Specifically, Morrow's account of *how* the hazard happens is not sensitive to potential variations in *who* the relevant agents are. Hale's account is silent on variations regarding both *who* the relevant agents are and *how* the hazard occurs. Corner and Pidgeon have a narrow understanding of *how* the hazard may happen and also lack

[106] The general concern about moral hazards is one whose history can be traced back at least to the Victorian era (Baker, 1996).

sophistication in their understanding of *who* the agents could be. For example, whilst Corner and Pidgeon identify that the hazard could happen at the political level, they fail to delineate the different relevant actors at this level, which is one of the tasks of this chapter.

The main contribution of this chapter is to provide a degree of conceptual clarity about the moral hazard which is absent from the literature. The five-variable understanding of a complete moral hazard complaint provides a helpful way of looking at the moral hazard literature and for formulating moral hazard complaints. This will be seen in section 4.6 below, where a table is generated based on how authors and reports seem to answer these five variables. This enables an analysis of moral hazard accounts to be undertaken, which helps point to strong and weak accounts as well as pointing to general limitations in the literature.

There is also an important question about the role which empirical evidence should play in determining whether we should respond to the moral hazard problem. This chapter observes that empirical evidence has been used in problematic ways by the likes of Reynolds (2015), who use it to cast doubt on the existence of the moral hazard problem. It then proceeds to provide an account of the role empirical evidence plays by drawing on the work of Shue (2010) around threshold likelihoods. The application of this approach provides an alternative account of the role of empirical evidence; one which, this chapter will argue, is more appropriate.

To summarise: this chapter seeks to make two contributions to understanding how the moral hazard problem is used in the context of SSI. The chapter provides a clearer conceptualisation of the hazard by identifying and considering the five variables which a complete hazard complaint requires. Secondly, it justifies taking this account of the hazard seriously, despite the inconclusive nature of the empirical evidence which is available.

4.1.1 The moral hazard complaint in governance reports: vague and ambiguous

The moral hazard concern is one which has a noteworthy place in public discourse on geoengineering ethics. A substantial number of geoengineering governance reports at least acknowledge the concern (Shepherd *et al.*, 2009; the SRMGI report, 2010; Climate Intervention: Reflecting Sunlight to Cool Earth, 2015 Schäfer *et al*, 2015; and the report of the Academic Working Group on Climate Engineering Governance ‘Governing Solar Radiation

Management' by Chhetri *et al.*, 2018). Many of these reports also appear to take the moral hazard concern seriously, as shown by the following examples.

The Royal Society:

> The very discussion of geoengineering is controversial in some quarters because of a concern that it may weaken conventional mitigation efforts, or be seen as a 'get out of jail free' card by policy-makers (Submission: Greenpeace; Submission: IOP; Submission: Lewis-Brown). This is referred to as the 'moral hazard' argument. (Shepherd *et al.*, 2009: 37)

The Solar Radiation Management Governance Initiative (SRMGI):

> Research into SRM could present a 'moral hazard'. If people (or governments) feel that they could be protected against the potential consequences of climate change, they may be less likely to take the actions necessary to reduce greenhouse gas emissions. In that case emissions would continue and rise (probably at a faster rate) and conceivably increase the eventual desire to deploy SRM technology. (SRMGI, 2010: 20)

And the National Academy of Science (NAS):

> Early discussions on albedo modification research focused on the so-called 'double moral hazard' issue – that on the one hand research into these proposed techniques could lead to policy-makers deciding to lose focus and/or urgency for reducing emissions. (NAS report, 2015: 152)[107]

These reports capture the idea that the moral hazard represents a concern about how SSI implementation could have an undesirable impact on the approach to mitigation. Yet the concept of the hazard is under-specified in these reports, and insofar as they do specify what it means, they fill out the idea in divergent ways. There is divergence about the stage of SSI at which the hazard occurs, for example. The Royal Society report points to 'discussions of geoengineering' as being enough for the hazard to occur, without specifying which discussions matter or by whom (Shepherd *et al.*, 2009: 37) . Whilst the NAS and SRMGI reports identify the hazard as occurring when research into SSI takes place, they diverge on the question of who the relevant agents are. The NAS report focuses on 'policy-makers' makers' (NAS report,

[107] The quote proceeds by explaining an alternative type of moral hazard argument where there is a situation in which SSI would need to be deployed but has not been properly researched.

2015: 152) whilst the SRMGI report focuses on 'people (or governments)' (SRMGI, 2010: 20). Nonetheless, they are more specific than the Royal Society report, which does not specify which agents it has in mind when considering the moral hazard problem. This level of vagueness about who the relevant agents are is deeply problematic if one wishes to address the moral hazard problem. For example, a response to the hazard which sought to regulate the behaviour of individuals may be inappropriate if the relevant agents are multinational corporations.

These instances of variation and vagueness across reports are symptomatic of a lack of conceptual clarity about what the hazard actually is. However, despite this lack of conceptual clarity, there is a consistent core claim across these reports, which is that SSI could have an undesirable impact on the process of mitigation. As explained, to achieve conceptual clarity, this chapter will answer the questions of what, who and how, which these reports fail to answer.

4.1.2 A comment on the wider moral hazard literature.

Before exploring how the moral hazard concern should be understood in the case of SSI, it is important to clarify the approach of this chapter. In particular, it is helpful to explain why the chapter does not draw more extensively on the broader literature about moral hazards which exist throughout the economics literature, on issues ranging from car (Weisburd, 2015) and health insurance (Einav *et al*, 2013), to the behaviour of banks (Boissay *et al*, 2016). This classic presentation of the hazard in the economics literature understands the hazard as occurring when there are two parties, an insurer and the insuree, and the insuree takes more risk than they otherwise would have done in light of being insured against the full consequences of that risk. There are two classes of objections to the insurance frame which are relevant in this case. As explained in chapter 2, Fragnière and Gardiner provide a normative argument against the insurance frame. Among other points, they argue that the insurance frame makes SSI look like a legitimate and desirable response to climate change, and by doing so hides the causes of concern with SSI (Fragnière and Gardiner, 2016: 16-17). Yet there is also an argument against the insurance frame which is more descriptive in nature; the argument is that the insurance frame simply does not fit the case of SSI. The difficulty of understanding the moral hazard of SSI in terms of the insurer-insuree relationship becomes clear when we seek to identify the parties who would occupy each of these roles.

Here I shall explore the idea that that insurance frame simply does not fit, by pointing to the significant conceptual obstacles one would need to overcome if SSI were to be described as a case of insurance. To demonstrate these obstacles, I will outline some of the work by Jonathan Wolff (2011, 2019) on risk, which demonstrates the varying structure of decisions about risk involving multiple parties, one of these situations being when the moral hazard exists. Wolff seeks to provide a framework for understanding situations in which risks exist which cannot appropriately be analysed via cost-benefit analysis in a public policy decision.[108]

	Decision Maker	Benefits go to	Costs go to
Individualism	A	A	A
Paternalism	A	B	B
Moral Hazard	A	A	B
Moral Sacrifice	A	B	A
Adjudication	A	B	C

(Wolff, 2019:10)

I shall only focus on Wolff's understanding of the moral hazard at the moment, as that is all that we require to have a sense of the technical understanding of the hazard. The other situations of risk which this table identifies are considered in chapter 5 as possible responses to the hazard. Wolff's understanding of the hazard is representative of the technical account of the moral hazard. The first feature to note of this account is that there are two distinct parties, in this case party A and B. Secondly, these parties have different roles in this case. One of them will have decision making power over the risk and will benefit from any risk-seeking behaviour, whilst the other party will bear the burdens of the risk-taking. This is a simple and fair account of the hazard, yet consider the following obstacles and barriers which exist when applying this account to the case of SSI.

There is a deep problem of understanding who the provider of the insurance is, and who the insured parties are. Are the insured parties those who are vulnerable to the effects of climate

[108] Recall the argument from part 2.2 of this thesis about cost-benefit analysis, where the point was made that it is inappropriate for many decisions about SSI due to it hiding ethically important considerations.

change which SSI seeks to prevent? This may make sense at first glance, but it is tainted by the fact that those 'insured parties' are likely to be vulnerable to some of the consequences of SSI. Does it make sense to say that you are meaningfully insured against a significant risk via a tool which would impose a new set of significant risks upon you instead? Moreover, if the hazardous behaviour is a product of certain actors not mitigating, then these non-mitigating parties are the ones who experience an increase in their risk-taking behaviour. So perhaps it is those parties who change their mitigation behaviour who benefit from SSI, so are they party A? Yet if this is true it contradicts the idea of SSI being insurance for the vulnerable against the harms of climate change, because it is not the vulnerable actors who we think of as experiencing this change in their mitigation behaviour. We think of those agents who already have high carbon emissions as those whose behaviour might change as a result of SSI.

These questions are indicative of the challenges one would face if they wished to engage with classic literature on moral hazard, which is dependent on there being the insurer and insuree relationship. This is theoretically challenging and possible answers to these questions appear to be convoluted, for SSI does not appear to have the simple insurer-insuree relationship. Yet the core claim of the hazard in the case of SSI is not contingent upon such a relationship. This is recognised in the work of Baatz (2016), Hale (2012), Morrow (2014) and Reynolds (2015), none of whom draw deeply on the wider moral hazard literature, when considering the moral hazard of SSI. In fact, despite their constructive engagement with the moral hazard in the case of SSI, none of them answer the above questions about the insurer insuree relationship. Hence this chapter is following a trend in the literature on SSI and the moral hazard, by not engaging deeply with the broader moral hazard literature.[109]

The challenges of adopting the insurance frame provides a reason to adopt a different frame which does not have these challenges. The risk-risk trade off frame is immune to the challenges the insurance frame faces, for nothing about the risk-risk trade off frame requires the A- B relationship between agents, and the subsequent challenges this creates. Yet the risk-risk trade-off is also informative of how we should think about the moral hazard problem. The moral hazard concern rightly points to the risk of there being an adverse impact on our mitigation

[109] One insight to draw attention to from the classic moral hazard literature is that this is an ex ante hazard, which is to say the hazard occurs prior to the event (such as the use of SSI). Whilst there may be an ex post hazard to be had from SSI, that is not the focus of this thesis, or any ethical analysis of the moral hazard from SSI so far. The ex post hazard would be a hazard which occurs after the event. If the event is the use of SSI, then the ex post hazard would be a concern about how our mitigation behaviour is affected

behaviour due to our research and development of SSI.[110] However, the moral hazard problem does not invite consideration of the danger of taking the moral hazard seriously. Yet the risk-risk trade off frame requires us to ask that question: What are the risks of taking the moral hazard seriously? The answer to this question is contingent upon how we would respond to the hazard. If our responses to the hazard restrict our knowledge of, or capacity to research or ability to deploy SSI, it would adversely impact the possibility of SSI being used to respond to certain risks of climate change.[111] Therefore responses to the moral hazard problem of SSI may increase the likelihood of certain risks from climate change being experienced.[112] Hence the importance of the risk-risk trade-off frame, which vividly reminds us of the importance of asking what the risks are of proposed responses to the hazard.

4.1.3 The difference between the moral hazard concern and the slippery slope concern.

It is also helpful to clarify how the moral hazard concern relates to the slippery slope objection, which is sometimes raised in the context of SSI. In the context of geoengineering the slippery slope concern is one of how initial action on geoengineering (which lies at the top of the proverbial slope), such as creating mechanisms of governance or doing lab-simulations, puts pressure on us to take further steps to advance geoengineering. This pressure could occur via some form of vested interested in SSI. For example by scientists who careers are progressing via this research, or by businesses and governments who have invested in the initial research. This pressure could lead small scale research to progress to large scale field tests and ultimately resulting in the deployment of SSI (which is at the bottom of this slope) (Callies, 2019b: 676). This is distinct from the moral hazard concern, which is a concern about how our mitigation

[110] This understanding of the hazard is explained in more detail in the following section of this chapter.

[111] These are plausible ways to address the hazard problem, chapter 5 argues that restricting knowledge via secrecy is an effective (but not desirable) way to address the hazard. The thought behind restricting research or deployment capabilities is that this should reduce the confidence which an agent has in SSI to be used to address risks from climate change.

[112] However, the picture is even more complicated than this, because these possible restrictions on SSI are designed to avoid the effect of the moral hazard. Therefore these restrictions on SSI subsequently restrains some risks from climate change (via the change in mitigation efforts) which SSI would seek to prevent. Yet the insight of using the risk-risk trade off frame still remains, which is when considering responses to the moral hazard problem we should consider how these responses interact with risks from climate change.

behaviour is affected by geoengineering. Even if both of these concerns ultimately result in the deployment of SSI the mechanisms by which this happens are different.

4.2 Why would the use of SSI technology pose a moral hazard?

In this section I will explore how moral hazards should be understood in the context of SSI according to Ben Hale (2012). Hale explains that there is ambiguity regarding how the hazard should be formulated (Hale, 2012: 120). This will therefore provide a helpful foundation for understanding what the hazard is in the case of SSI (although, as Hale acknowledges, the hazard is still not completely clear even after these questions have been answered).[113] This section is therefore a starting point to show what key parts of the literature on the ethics of SSI have to offer in terms of understanding the hazard, and it provides a springboard for the further inquiry into understanding the specific hazard which this chapter does address. Therefore, after an initial engagement with Hale's work, the chapter will move on to engage with further components of the hazard for the purposes of our analysis.

If we take SSI technology seriously as a way of engaging with climate change, are we more likely to end up in a situation in which it will be required? If you think the answer to this question is 'yes', or you can see why one might think it is 'yes', then you can see the intuition behind the concern that SSI creates a moral hazard. Yet this is too ambiguous for the purpose of our analysis. Hale illustrates the ambiguity of the moral hazard by observing that the hazard can refer to at least three different objections in relation to different emission pathways.[114] These are the business-as-usual, the counterfactual trajectory, and the perverse behaviour objections. The three objections can be represented as shown in Table 2.

Table 2. Three Different Outcomes of the Moral Hazard

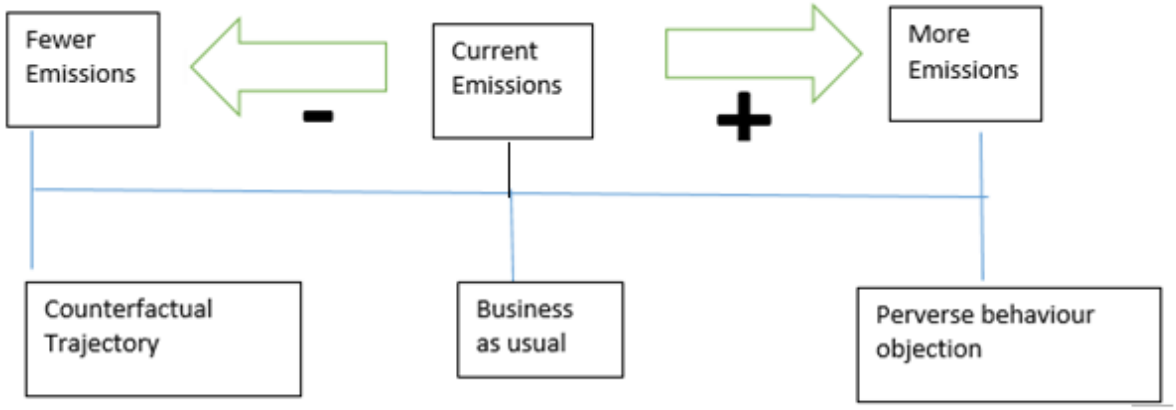
The business-as-usual objection	There will be no change in our collective mitigation behaviour.
The counterfactual trajectory objection	Whilst mitigation efforts will increase, they will not increase by as great an extent as previously thought.
The perverse behaviour objection	We will mitigate less than we currently intend to.

The table is based on Hale's understanding of these objections (Hale, 2012:120–122).

[113] To be more specific, this means that at this stage of the chapter certain features of the hazard will remain ambiguous, such as who the relevant agents are or how the hazard occurs.
[114] Hale's account relates to geoengineering technologies in general, not specifically SSI.

The concerns identified by Hale indicate that the hazard can be understood in significantly different ways. Despite this, these three concerns are all of the same metric: namely, emission levels. The fact that these concerns are of the same metric allows us to understand them as existing on a spectrum. Figure 3 contains a diagram which I have created to illustrate this point.

Figure 3. The different types of Moral Hazard concerns.



4.3 When would SSI create a moral hazard?

Despite the clarity gained by understanding the moral hazard complaint in terms of SSI implementation having an adverse impact on emission levels, there is still some ambiguity in Hale's formulation of the moral hazard complaint. This is because it is not clear when SSI can have this hazardous effect. Is it when agents have knowledge of the concept of SSI? Or once it has been researched and developed? Or when it is ready to be deployed? Or at some other stage in the development of SSI? There is a range of points in the process of SSI development where one may fear that the moral hazard exists.

In order to clarify the nature of the moral hazard concern which this thesis will analyse, I will focus on the concern when it is expressed in relation to the research and development of SSI. Moreover, by focusing on the research and development stage of SSI, the thesis retains its relevance for governance reports, many of which express concerns with the moral hazard at the stage of research and development:

> Research into SRM could present a 'moral hazard'. (SRMGI, 2010: 20)
>
> There is a risk that research on albedo modification could distract from efforts to mitigate greenhouse gas emissions. (NAS, 2015: 152)
>
> [T]he 'moral hazard' argument (the concern that research on climate engineering would discourage the overall efforts to reduce or avoid emissions of greenhouse gases). (Schäfer *et al*, 2015: 14)

An additional reason for focusing on the research and development stage is the fact that this stage may occur soon, given the moves which are being made for experimental research on certain aspects of SSI to take place (Tollefson, 2018).[115] The relevance of focusing on the research and deployment stage of SSI is supported by the thought that agents are more likely to take an idea seriously if it is being researched and developed, particularly if the research supports the possibility of SSI. Therefore this chapter will proceed by understanding the moral hazard complaint as the concern that the research and development of SSI may have an adverse impact on mitigation levels.

4.4 Why is the moral hazard bad?

The final variable to consider in this section is why the moral hazard may be considered to be bad. This is a valuable exercise, for if the hazard was not bad in some sense, then it would be unclear why we should be taking it seriously and why it is being taken seriously in public policy domains. This chapter will understand the badness of the hazard as a product of how it increases the likelihood that we will face risks from climate change, the use of SSI, or both. The idea here is that an adverse impact on mitigation efforts will lead to a scenario in which there is a greater degree of climate change and/or a scenario in which SSI is more likely to be used because there is a greater risk of climate change. This seems to be an implicit assumption in much of the literature, given that much of it is written in the context of debates about climate justice or producing reports which hold climate change to be a problem.

[115] If one counts computer simulations as part of the research and development stage, then research is already happening.

4.5 Who are the relevant agents?

This next section considers the question of whose behaviour changes in order for an adverse impact on mitigation efforts to occur. There has been an attempt in the literature on moral hazard and geoengineering to identify the relevant agents. Corner and Pidgeon (2014) identify three relevant levels where the hazard could occur. Firstly, where an individual adversely changes their mitigation efforts (Corner and Pigeon, 2014: 2). Secondly an individual may not change their mitigation behaviour in direct response to the prospect of SSI, but in response to the fear that others in society will do so, then society as a whole may change its attitude towards mitigation as a consequence. Thirdly, there is a political hazard concerning where climate policy is formed and how the resources at hand could be moved away from mitigation efforts (Corner and Pigeon, 2014: 3).

If we are going to understand who the relevant agents are, we will need to build on Corner and Pidgeon's trichotomous account of individual, societal and political actors. I am going to focus on the political actors; a reason for doing this is that political actors are the appropriate agents to be governed by a global governance institution.[116]This is because if we were to accept their analysis and focus on the political level, there would still be an absence of necessary detail, for it would still not be clear who the relevant actors are in climate policy at the political level. To answer the question of who the relevant actors are at the political level I will consider what a polycentric approach to climate policy formulation tells us. This section of the chapter will therefore identify a sub-set of relevant agents at the political level, namely: the formal role-holders of power. The formal role-holders of power are understood as the relevant agents of the moral hazard complaint for this thesis. By the formal role-holders of power, I mean those agents who are members of the three branches of government within a state: the legislature, the executive and the judiciary.[117] This is done in the knowledge that there is a certain tension, which is that polycentrism, in part, encourages us to think beyond the state in terms of how policy formation occurs. However, it does not deny the importance of the state, it simply points

[116] There would be a distinct strangeness to individuals being directly governed by a global governance institution.
[117] The relationship between these branches will of course vary between states. If we were to consider the case of the United States of America the total number of formal role holders of power is 545, the number 545 comes from the 535 members of Congress plus the 9 members of the Supreme Court plus the President. The number is higher if the Cabinet is also included, but the number of Cabinet members can vary within each administration. The formal role holders of power should be identifiable in each state, by summing the members of the legislature and executive ad judiciary, although if one were to engage in this exercise there would be a question to be asked about double counting where there is overlap between these branches, for example in the UK members of the legislature are also members of the executive.

out other relevant agents alongside the state. The point of drawing attention to polycentrism is that it reveals who the possible relevant political agents are for the purpose of mapping different accounts of the moral hazard.

The descriptive claim of polycentric governance is that climate policy is made at many different levels of governance ranging from international fora such as the UNFCCC to local councils (Jordan *et al.*, 2018). Moreover, these policies are in part a product of the perceived interests of many groups, including the government, the people, local governments, businesses and non-governmental organisations (NGOs). For example, NGOs and businesses are capable of independently pursuing policies and norms regarding climate change within their domain. Consequently, this would have a degree of influence on climate policy. The thought is that climate policy is not only the product of monocentric action by states at the international level; it is also the product of complicated relationships between states and other agents.

Polycentrism helps make sense of the importance of actors who are outside the government in affecting climate-related polices. These are all those actors who could have an adverse impact on mitigation if they were to experience the moral hazard.[118] Yet it is not necessary for all of these agents to be subject to the moral hazard for the unwelcome consequences of the hazard to materialise. Instead, there is likely to be a threshold number of relevant agents who need to experience it in order for the hazard to have an adverse impact on mitigation.[119] This variety of actors who are relevant to mitigation policy and therefore also susceptible to the hazard points to a difficulty with trying to understand how SSI could create a moral hazard.[120] Different types of agents might experience the hazard under different circumstances and due to different reasons. For each type of agent there is a story to tell about the conditions which would lead them to be susceptible to the hazard.

Although this threshold image might be misleading, for not all agents wield equal power in being able to influence the likelihood of the moral hazard that occur in public policy. Public policy dynamics are complicated and are not going to be explained here. But the point is a simple one, that even if many agents who do influence public policy do not experience the

[118] Their impact on mitigation could occur in a variety of ways, ranging from reducing their personal emissions to their ability to influence others, and they could affect decisions at the regional, national or global level.
[119] Even if we do not know what this number is, it seems unlikely that all agents would have to be subject to the hazard in order for the hazard to have an adverse impact on mitigation.
[120] The hazard can of course also be experienced by those outside the policy process. For example, individuals could change their own behaviour in light of SSI.

moral hazard, this may be insufficient to prevent the moral hazard occurring if particular agents who inform public policy do experience the hazard.

For the purposes of this thesis, I will focus on one of these relevant groups of actors in the knowledge that my subsequent analysis only addresses the moral hazard problem for those actors, and not the question of the hazard as a whole. I will therefore focus on the formal role-holders of power as the relevant agents in my analysis of the moral hazard of SSI.[121]

The relevance of the formal role-holders of state power can be seen when one considers why the state itself is such an important agent in climate change policy. The importance of states in a polycentric account of climate change policy is vividly made by Nachmany and Setzer (2018). They make the point that states are responsible for formulating and implementing nationally determined contributions (NDCs) as set out in the Paris Agreement (Nachmany and Setzer, 2018: 48). The NDCs are an integral part of the current strategy to address climate change, and this strategy relies on states playing their part and acting to meet their NDCs.

Crucially, the state has the power to regulate the behaviour of other actors and to encourage them to behave in certain ways (Nachmany and Setzer, 2018: 49). The state wields its power in a unique way, and in regard to the power to regulate behaviour, no other institution has as much legal force to regulate the behaviour of agents within a state's territory. Additionally, the state is uniquely placed to motivate behaviour at home – and to some extent abroad – given the range of incentives it has to offer, ranging from tax breaks to favourable laws. This combination of power to regulate and incentivise behaviour leads Nachmany and Setzer to conclude that the state has a 'central role' in climate governance which cannot be performed by non-state actors. These powers of the state are indicative of the state being the primary institution with governing power in regard to climate change.

The powers of the state that Nachmany and Setzer (2018) identify are powers which are formally exercised by the formal role-holders in the state. For example, the power to regulate behaviour is formally a product of the relationship between the executive, legislature and judiciary. Given the importance of these powers, I take those agents who formally exercise them as offering a sufficiently interesting level of agency to focus on for the analysis of the moral hazard problem in the case of SSI. It is important to emphasise the word *formal* role-

[121] This is done in the knowledge that there is a certain tension, which is that polycentrism in part encourages us to think beyond the state in terms of how policy formation occurs. However, it does not deny the importance of the state, it simply points out other relevant agents alongside the state.

holders here. Members of Parliament (MPs) can be seen as the formal role-holders of legislative power, yet there could also be agents outside the parliamentary process who are capable of exerting great influence on the decisions which are made in Parliament. These other agents are not the unit of analysis for this thesis.

To summarise, the preceding discussion was motivated by the premise that a complete formulation of the moral hazard account requires an account of who the relevant agents are in the moral hazard concern. Polycentric accounts of climate policy formulation draw attention to the fact that there are a range of relevant agents involved in climate change policy formation. Nachmany and Setzer argue that states are one such relevant agent due to the regulatory and incentivising power which they have. Given Nachmany and Setzer's insights about the importance of the state's powers, the chapter proceeds by taking those who are in a positon to formally exercise these powers as the relevant agents for our account of the moral hazard concern.

4.6 What mechanisms may account for these agents to experience the moral hazard?

In this section I shall consider what mechanisms could lead the formal role-holders of power to experience the moral hazard when SSI is being researched and developed. Whilst this section focuses on the formal-role holders of power there is nothing about these mechanisms which is unique to the formal role-holders of power, therefore this section also continues the mapping exercise of thinking about what mechanisms could lead to agents in general experiencing the moral hazard.[122] The thought is that by understanding these mechanisms we will be better placed to identify ways to address the hazard. Morrow (2014) identifies three mechanisms which will be informative for the analysis. These are: informational failure, psychological failure, and ethical failure.[123] This section will therefore proceed by exploring each of these mechanisms.

4.6.1 Informational failure

Broadly speaking, informational failure occurs when the relevant agent fails to understand or have access to all the relevant information they need to make a decision. The informational failure which Morrow refers to is grounded in economic models (Morrow, 2014: 7). Economic models of a certain kind tend to assume that an agent has full information of a risk when making their decision. In short, the informational failure would occur if the policy-maker is unaware of the probability of all possibilities and the possible effects from each possible policy (Morrow, 2014: 7). Morrow rightly observes that the assumption of full information does not necessarily hold in the case of climate engineering policy-makers (Morrow, 2014: 7). This is hardly surprising, given that in the case of SSI there is a huge amount of information to be understood if one is to truly understand the outcome and magnitude of each possibility. This

[122] Although different agents may be effected by these mechanisms to different degrees.

[123] A limitation of Morrow's analysis is that he focuses on policy-makers without defining the term. Morrow does not appear to be sensitive to the insights of polycentrism, i.e. the fact that there are other agents who may experience the moral hazard and that their experience of the hazard is relevant for mitigation efforts. Or, if Morrow does capture all these agents in his understanding of policy-makers, then his analysis suffers from not being sensitive to how the responses of different agents would vary in line with these mechanisms.

sub-section will proceed by elaborating on what an informational failure is in the case of SSI and why such a failure may contribute to the moral hazard.

It is important to note that the informational failure is a descriptive claim in context, not a normative claim. Whether an agent is blameworthy, responsible or should be punished is a separate set of questions. Morrow identifies an informational failure as occurring in the case of climate engineering when policy-makers 'remain ignorant or even develop false beliefs' about the 'promise, efficacy and risks of climate engineering technologies' (Morrow, 2014: 7). Therefore, the descriptive claim is that if the formal role-holders of power remain ignorant, or have developed false beliefs about the promise, efficacy and risks of SSI, then there is an informational failure.

It appears likely that informational failure will be quite pervasive among the formal role-holders of power in the case of SSI. The question to ask is how particular instances would contribute to the moral hazard happening? The concern is that the informational failure could contribute to the hazard due to the formal role-holders of power perceiving SSI as being more effective and risk-free than it actually is. For example, if a formal role-holder of power is ignorant of the risks of SSI, they may perceive it as a sufficient response to climate change. Such an attitude can be seen in the blog post of Newt Gingrich (Former Speaker of the United States House of Representatives) (Gingrich, 2008).[124] Yet not all informational failure about SSI would contribute to the moral hazard. For example, if an informational failure results in one only knowing about the risks of SSI but none or few of the benefits of SSI such as preventing or reducing temperature increases, this may result in the agents having an unduly *un*favourable view of SSI. Our concern is whether through ignorance or false belief the formal role-holders of power will have an unduly *favourable* view of SSI, since this may have an adverse impact on mitigation efforts, which is the cause of our fear about the moral hazard.

For each of the three mechanisms which contribute to the existence of the moral hazard it is useful to think about why these mechanisms may occur. We therefore need to ask why might an informational failure occur which could contribute to the hazard? Here I shall focus on ignorance rather than the development of false belief, a reason for this focus is that at this point in time ignorance about SSI is common. Survey data in the US revealed that 74% of the general population knew nothing about geoengineering (Leiserowitz *et al*, 2010). A similar degree of ignorance has been shown in the UK where 75% of respondents to a study claimed to know

[124] Gingrich's position has been explained in section 1.2.1. of this thesis.

nothing or almost nothing about geoengineering (Spence *et al*, 2010: 20) It is reasonable to extrapolate that many formal role holders of power would also be ignorant of geoengineering.

When considering ignorance about SSI, it is important to draw attention to the fact that this is the default setting for many people. Formal role-holders of power may educate themselves about SSI, listen to experts, read reports, meet advisers, etc. in order to overcome this ignorance, but the default state is one of ignorance about the risks, efficacy and promise of SSI. Even when a formal role-holder does educate themselves about SSI, there is no guarantee that this would be to a sufficiently high standard to escape the state of informational failure. This is not a claim about the intelligence of formal role-holders of power; it is simply pointing to the fact that SSI has many technical, social, ethical, political and scientific implications which are challenging for many people to understand. What I mean is that we do not need to look for a deep complicated theory to account for the ignorance of the formal role-holders of power on this issue, since ignorance to a degree is natural and to be expected in this case.

To summarise, an informational failure occurs when the formal role-holders of power are ignorant or have false beliefs about the promise, risks and efficacy of SSI technology. Informational failure which leads to formal role-holders of power having an unduly favourable perception of SSI could result in the moral hazard occurring. We have reason to fear that an informational failure could be pervasive among the formal role-holders of power due to the fact that ignorance of the risks, promise and efficacy of SSI is the default state of knowledge for many role-holders. A more complete account of informational failure in relation to SSI would also consider how those who formally hold power come to develop false beliefs.

4.6.2 Psychological failure

The second mechanism which Morrow identifies is that of psychological failure, or cognitive bias (Morrow, 2014: 7). Cognitive biases are ‘unconscious psychological mechanisms’ which skew a judgement (Morrow, 2014: 7). Judgements about SSI are judgements about risk (broadly speaking).[125] Kahan *et al.* (2015) present three conflicting models of risk perception. These are the rational actor model, the irrational actor model, and the cultural cognition thesis. This sub-section will show why we have reason to suspect that SSI proposals could interact

[125] This is clear if we remember the risk-risk trade-off frame, moreover this view is prominent in other works on the moral hazard, such as Baatz (2016) and Morrow (2014).

with the irrational actor or cultural cognition accounts of risk and so contribute to agents operating in the manner predicted by the moral hazard concern.[126] By doing so this sub-section will show that we have reason to suspect that SSI may interact with well-established cognitive biases.

The cultural cognition thesis has a story to tell in this case. The thesis states that when evaluating risk, the valuation will be skewed by that agent's own values. Specifically, it is skewed so that the *risk* coheres with the agent's own values (Kahan *et al.,* 2015: 194). Kahan *et al*. provide the example of an agent with individualistic values[127] who would oppose a response to climate change which would curtail their liberty (Kahan *et al.,* 2015: 194). The relevance of the cultural cognition thesis to the SSI case seems to be dependent on whether the risks are ones which interact with an agent's value judgement as to how society should function. That would definitely appear to be the case, since the risks from SSI are physical, social, economic and political. Albert Lin draws attention to the fact that SSI interacts with how people value 'human innovation, faith in technology and the dominance of nature' (Lin, 2013: 700).

The moral hazard could therefore occur if people hold values which result in them being unduly optimistic or trusting in the capacity of SSI to be a good response to the problem. Consider the example of the techno-optimist who holds the view that technology will solve the problem of climate change. Upon learning about SSI, the cultural cognition thesis would expect the techno-optimist to look favourably upon SSI because it is consistent with their image of how society should be. Consequently, they may reduce their support for mitigation efforts because they believe that SSI is the technology which will solve the problem. Hence they behave in a way which would appear conducive for the moral hazard to occur.

I shall now identify an alternative account of how the mechanism of psychological failure could lead to the moral hazard, namely the irrational actor model (Kahan *et al*., 2015: 195). This model is premised on the idea that an agent's perception of risk is affected by a set of heuristics and biases (Lin, 2013: 694). This model is well established in the psychological literature (Slovic, 2000; Kahneman, 2003; Weber, *et al* 2006) and has started to infiltrate other academic

[126] There are two reasons for not paying much attention to the rational actor model. Firstly, the idea of all actors being rational utility maximisers is outdated and the cultural cognition thesis and the irrational actor model both appear to be more plausible accounts of human psychology. Secondly, the model would say very little of value in this case, apart from that agents will engage in behaviour which contributes to the moral hazard when it is in their perceived rational self-interest to do so.

[127] For Kahan *et al*. this means they value having a broad sphere of liberty (Kahan *et al*., 2015: 194).

disciplines as well. Recently the Nobel Prize for Economics was awarded to Richard Thaler for his work in the area of behavioural economics, which is premised on this irrational actor model (Thaler, 1991; 1994; Thaler and Sunstein, 2008).

When the irrational actor model is applied to a formal role-holder of power who is making a judgement about SSI, the theory would seek to identify whether the agent is making a judgement about risk. Each judgement about risk opens a window for heuristics and biases to come into play. In the case of SSI there are plenty of risks involved, ranging from what one thinks of the physical impacts of SSI to the political ramifications of taking SSI seriously. Hence there are plenty of opportunities for heuristics and biases to come into play as well. The question is whether we have any reason to suspect that certain heuristics or biases would cause the formal role-holders of power to contribute to the moral hazard?

Lin identifies quite a few relevant heuristics and biases, namely: optimism bias, overconfidence, hyperbolic discounting, and outrage (Lin, 2013: 695). Lin explains that the optimism bias is when people do not believe they will suffer the consequences of a phenomenon like climate change if there is any uncertainty about whether harm will occur or not (Lin, 2013: 696).[128] The optimism bias could manifest in the case of SSI. An instance of this would be if the formal role-holders of power were aware that there is a set of risks associated with SSI, yet believed that these risks are more likely to be experienced by other countries than their own.[129] The formal role-holders of power who experience this overconfidence may be more willing for SSI to be adopted as a response to climate change, given that they do not think their own country is at risk. Consequently, this may have an adverse impact on the willingness of those formal role-holders to engage with demanding efforts for climate change mitigation. This is merely meant to illustrate why we may have reason to suspect that heuristics and biases could be part of the causal mechanism for the moral hazard.

It is important to remember that judgements about SSI are not occurring in a vacuum, whether it is one's values, heuristics or biases which are informative of how SSI is perceived; they are all operating against the background conditions of the perfect moral storm. Consider an agent

[128] The idea of an optimism bias is well supported in the psychology literature (see Sharot, 2011; Sharot *et al*., 2007; and Weinstein, 1980). Moreover, the idea that individuals experience an optimism bias in the case of climate change is well explored (Nicholls, 1999; Hatfeld 2001; Gifford, 2011; Van Der Linden, 2015).

[129] There is a challenge in regard to identifying the optimism bias in the case of SSI, since the risks of SSI are not distributed evenly. Therefore it may well be statistically accurate for the formal role-holders of power to think that their country is less likely to be subject to the risks of SSI than others.

who has the value of technology optimism: it becomes even more appealing to place weight on that value if they consider the demandingness of the alternative strategies to navigate the perfect moral storm. Consider a heuristics or bias, such as overconfidence: it is one thing to be overconfident in the promises of SSI, it is perhaps even easier to be overconfident in the promise of SSI if you are an agent who does not bear the significant risks of SSI failing. When an agent has the values, biases or heuristics which we would expect to lead to the moral hazard, then we need to understand that they are occurring in the context of the perfect moral storm, which could exert an influence on them and even compound the force of these values, biases and heuristics.

This sub-section has considered two well-established theories of risk perception and illustrated that according to these theories we could make sense of how the moral hazard could occur among certain agents. This is not proof that the moral hazard would necessarily occur due to the psychological mechanisms in play; it is just to show that some of our best theories about responses to risk lead us to expect the moral hazard to occur. If we do want a greater understanding of how psychological mechanisms could contribute to different agents acting on the moral hazard, it would beneficial for there to be psychological research specifically focused on this question. Such research appears to be desirable if we do want to design effective mechanisms to prevent the moral hazard.

4.6.3 Ethical failure

The third mechanism which Morrow identifies is that of ethical failure (Morrow, 2014: 8). In this context, Morrow understands an ethical failure as one where ‘politicians rely too heavily’ on SSI. Morrow understands that politicians may be tempted to do this if they think that SSI would involve fewer costs than mitigation. One might think that making a decision about SSI on the grounds of cost does not represent a case of moral failure, since relative costs are a legitimate consideration in public policy decisions. Whilst this is true, a decision on SSI should not be made solely on the grounds of cost. As explained in chapters 1 and 2, there are many ethical salient consideration when one makes a judgement about SSI, and cost is not the sole or even a particularly important consideration when weighing these considerations.

The question is whether we can understand why politicians may experience an ethical failure in this case. Ethical failure in the case of climate change policy has been well highlighted by

Jamieson (2014), and Shue (2014). Moreover, Gardiner has provided a detailed account of why this failure occurs; as discussed in sections 2.3 and 3.2, he understands such failure as a product of the perfect moral storm. The key points to remember are that there are at least three storms – the global, intergenerational and theoretical storms – which converge to challenge our ability (and that of the formal role-holders of power) to act in an ethical manner in response to climate change.

Gardiner applies the perfect moral storm to provide an analysis of some cases of geoengineering. A case which Gardiner has given particular attention to is the 'Arm the Future' argument, whereby research into geoengineering is justified on the grounds that it will allow future generations to create a 'less bad' state of affairs for themselves (Gardiner, 2011: Chapter 10). A key feature of Gardiner's analysis is that geoengineering policy is highly susceptible to moral corruption. As explained in Chapter 3, one of the dangers associated with moral corruption in the case of SSI is how it could give rise to a dialogue about SSI policy where SSI is pursued in the name of it being ethically required, but really this is done due to it being in the self-interest of powerful agents. We have reason to fear that an SSI policy which aligns with the self-interest of the formal role-holders of power is one which meets Morrow's standard of ethical failure, whereby politicians are too reliant on SSI as a solution. The idea here is that SSI policy is much less costly to politicians than mitigation efforts. And moral corruption will tempt them to justify to themselves, and to others, an excessive reliance on SSI because it relieves them of the (perceived to be) more costly burden of acting on mitigation.

It is helpful here briefly to review the relationship between the concepts of ethical failure, moral corruption and the perfect moral storm. This subsection has drawn the link between moral corruption and ethical failure, by observing that moral corruption can be understood as a specific instance of ethical failure. Chapter 2 argued that our judgements about climate change in general, and SSI in particular, occur against the background conditions of the perfect moral storm. These judgements include ones which we can understand as ethical failures. Therefore we have a reason to understand the problem of ethical failure, and of moral corruption more specifically, in the context of the perfect moral storm. By contextualising instances of ethical failure against the perfect moral storm, we are able to have clearer understanding of the circumstances of these failures. In the case of moral corruption we can see the structure of the perfect moral storm makes moral corruption particularly tempting. Trying to address moral corruption without awareness that the perfect moral storm compounds the temptation of agents would be foolhardy.

Moral corruption is not the only way that ethical failure could occur. The morally corrupt agent at least cares about trying to justify their actions as being ethically permissible or even required. Yet we could also have ethical failure brought about by agents who simply do not value acting ethically, or even appearing to act ethically. A more complete analysis of ethical failure would also take into account such agents who do not care about ethics. Despite this, we do now have an insight into how ethical failure could occur in the case of SSI, which is consistent with, and grounded in, the broader literature about the ethical failure of climate policy.

To conclude this sub-section, ethical failure represents a possible mechanism by which the moral hazard could occur. Moral corruption provides a helpful way for understanding how ethical failure in the case of SSI could happen. We have strong reason to suspect that agents are susceptible to moral corruption, and the likelihood of moral corruption is compounded by the background conditions of the perfect moral storm. The behaviour resulting from moral corruption meets Morrow's standard of ethical failure, given that politicians would be relying too heavily on SSI.

4.6.4 A summary of the causes of the moral hazard

I have used this section to try to make sense of the different mechanisms which may create a moral hazard for the formal role-holders of power. By doing so, the chapter has provided a level of clarity about the moral hazard concern which is absent from geoengineering governance reports; these reports fail to provide any account of how the moral hazard occurs. This section has adopted Morrow's account of the three relevant mechanisms which can induce the moral hazard. Morrow identified these as: informational failure, cognitive bias and ethical failure. This is not a complete explanation of how the moral hazard can occur, not least because the processes of cognitive bias and informational failure for the formal role-holders of power are not yet well understood in this case. There may also be other mechanisms involved which Morrow does not consider.

4.7 The moral hazard table

Thus far the chapter has identified five variables which need to be included for any account of the moral hazard to be complete. These are (1) the stage at which the moral hazard occurs; (2) who the relevant agents are; (3) the mechanisms by which the hazard occurs; (4) the impact of the moral hazard; and (5) why it is a bad thing.

This next section now seeks to illustrate the usefulness of the five-variable understanding of the moral hazard. To do this, I shall apply the five variables to test how the term has been used in the SSI literature. By test, I mean I will explore the accounts of the moral hazard which exist in the geoengineering literature to see whether they take each variable into consideration, and if so, how. The benefit of this is that it clearly articulates how the moral hazard complaint has been formulated by others. This allows us to engage in critically exercises, a simple and helpful exercise is to observe whether authors have a complete account of the hazard or not. Moreover we can then observe where this is deviation in the literature about how these variables.[130]

4.7.1 Applying the five-variable test

David Morrow in his 2014 paper seeks to understand why the moral hazard may be bad. He adopts the following understanding of the hazard:

> Roughly, their argument is as follows: researching climate engineering would induce policy-makers to mitigate less than they otherwise would. (Morrow 2014: 2)

Morrow is clear that he is interested in the research and development stage of geoengineering, which is his answer to the first variable (Morrow, 2014: 2). As for the second variable regarding who the relevant agents are, it is clear from the above quote that this is policy-makers. Yet at no point in his article does Morrow elaborate as to how ‘policy-makers’ should be understood – should it be those who formally exercise policy-making power, or those who substantively

[130] This table could also be helpful for a more complicated type of analysis, whereby one could seek to determine better and worse accounts of the hazard by critically comparing them.

do so? This lack of clarity means it would be hard to address Morrow's account of the hazard without further detail about how the 'who' variable should be understood.

As we saw in section 4.3, Morrow is very clear about the mechanisms by which the moral hazard occurs. He argues that the moral hazard comes about through informational, psychological and ethical failings (Morrow, 2014: 7–8). The fourth variable focuses on what the impact of the hazard is. Morrow clearly states that the impact will be less mitigation than there otherwise would be (Morrow, 2014: 4–5). The final variable asks why the hazard would be a bad thing. Morrow answers this by arguing that it would be bad in circumstances where it resulted in there being a worse state of affairs than would otherwise be the case (Morrow, 2014: 4–5). Morrow does nonetheless provide a quite well-rounded account of the moral hazard, with answers to all five variables in his paper.

Corner and Pidgeon (2014) try to explore whether the moral hazard might exist at the level of the general public. The stage of the hazard which they consider is that of contemplating the use of geoengineering. This can be seen by the fact that their paper explores how the general public reacts to the possibility of geoengineering (Corner and Pidgeon, 2014:7). The second variable is that of who the relevant agents are. For Corner and Pidgeon, this is the general public and those who act at the political level (Corner and Pidgeon, 2014: 2–3). Corner and Pidgeon make it clear that they understand the political level as being the actions of politicians (Corner and Pidgeon, 2014: 7). This gives scope for plenty of questions to be asked. For example, why are politicians the only relevant political actors at the political level for this account?

The third variable concerns the mechanisms which would lead to the moral hazard occurring. Corner and Pidgeon do not provide an account of any mechanisms, but they do shed light on a related question, in that they consider how people may respond differently to the hazard depending on the different values that they hold. This may account for variation in the effect of mechanisms which may cause the moral hazard. Nonetheless, Corner and Pidgeon's account of the moral hazard concern leaves a variable to be desired, given the absence of mechanisms from their account. They hold the common view that the moral hazard complaint is about the introduction of SSI having an adverse impact on emission levels (Corner and Pidgeon, 2014: 1). As for the final variable – why the moral hazard is bad – this does not seem to be considered in Corner and Pidgeon's paper. This makes sense, given that their paper is concerned about whether the moral hazard exists rather than normative questions such as whether the moral hazard is bad.

Albert Lin's paper seeks to understand whether geoengineering could present a moral hazard (Lin, 2013). Lin provides a clear expression of his understanding of the moral hazard concern thus:

> The moral hazard concern is that research and development in geoengineering may undermine public and political support for mitigation and adaptation, notwithstanding geoengineering's limitations. (Lin, 2013: 678)

Lin's answer to our first variable is clear; he is focusing on the research and development stage (Lin, 2013: 677). And he understands the 'who' as 'public and political support'. It becomes clear in the following pages that by political support Lin means the support of policy-makers. This is clearly seen in the last sentence of Lin's article:

> In assessing geoengineering options, policy-makers and the public must remain cognizant of the moral hazard danger and take steps to counter it. (Lin, 2013: 711–712)

As for the third variable, Lin is useful in providing an analysis of how the psychological mechanism could play a role in creating the moral hazard (Lin, 2013: 694–699). Lin understands the impact of the moral hazard as one of undermining mitigation efforts (Lin, 2013: 711). And he does appear to be concerned with the risks of both climate change and SSI which may result from the moral hazard (Lin, 2013: 678). Lin does provide quite a complete account of the moral hazard since it is possible to answer all five of our variables by reading his paper. Yet if his account is the one which we wish to address, then we require more information about the 'who' variable, because, as with the others, the who for Lin is couched in the vague terms of 'policy-makers', and there is significant scope for variation in interpreting who the policy-makers are.

There is a frequently cited letter by Martin Bunzl in which he is critical of arguments which are used against the research and development of geoengineering technologies. One such argument is the moral hazard argument (Bunzl, 2009). Bunzl identifies the stages where the moral hazard could occur as being research and development as well as implementation. For the second variable – who – it is very unclear who Bunzl has in mind. Likewise, he does not consider our third variable; he does not offer an account of how the moral hazard could happen. As for the fourth variable, Hale (2012: 121–122) argues that Bunzl holds the perverse behaviour view of what the hazard is, due to Bunzl stating that the moral hazard would lead to an increase in the output of greenhouse gases. Finally, Bunzl does not articulate why the moral

hazard would be bad. Bunzl's scepticism about the moral hazard argument is hardly surprising, given that his account of what the problem is appears to be very incomplete, missing out three of the key variables (the agents, the mechanisms and why it would be bad). Given the incompleteness of Bunzl's view, this would seem to be a poor account of the moral hazard concern.

The five-variable test can also be applied to assess how the moral hazard concern is expressed in governance reports. I shall therefore proceed by applying the test to three reports. The Royal Society report is one of the first reports to be written on geoengineering, and in this report the moral hazard concern is understood like this:

> In the context of geoengineering, the risk is that major efforts in geoengineering may lead to a reduction of effort in mitigation and/or adaptation because of a premature conviction that geoengineering has provided 'insurance' against climate change. (Shepherd *et al*., 2009: 37)

The Royal Society report is inconsistent as to what stage of geoengineering it thinks the moral hazard may occur at; it warns that the 'very discussion' may be enough for the moral hazard to occur (Shepherd *et al*., 2009: 37). Yet in the quote above the Royal Society clearly thinks that the moral hazard could occur if there are major efforts at geoengineering. 'Major efforts' could plausibly be understood as representing more than mere discussion, yet the report does not elaborate on what exactly it means. Therefore I understand the report to be imagining that the moral hazard occurs at the stage of contemplating SSI.[131] The report does not state who it thinks the relevant actors are, the mechanism by which the hazard could occur, or why the hazard would be bad. The only other variable the report is clear about is the impact of the moral hazard and that it leads to a reduction in mitigation and/or adaptation efforts (Shepherd *et al*., 2009: 37). Given that this report raises the moral hazard problem as one which ought to be addressed if the empirics show that it does exist, it is troubling that conceptions of the moral hazard are so incomplete. Also, if the moral hazard is to be addressed, then we need to know who the relevant agents are, the mechanism by which the hazard occurs, and why it is bad.

The EUTRACE report (Schäfer *et al*, 2015) seeks to provide a broader understanding of geoengineering's scientific, social and governance challenges (Schäfer *et al*, 2015: 13–15) and

[131] I get the language of 'contemplating' SSI from Preston's literature on the ethics of geoengineering, in which he deifies that concerns about geoengineering can occur at the stages of; contemplating geoengineering, research and developing geoengineering, deploying geoengineering, or post-implementation of geoengineering. (Preston, 2013)

it devotes at least three pages to considering the moral hazard problem (Schäfer *et al*, 2015: 58–60). The report presents its understanding of the moral hazard concern thus:

> A prominent concern around climate engineering has been the fear that discussing, researching, and developing climate engineering techniques may have negative effects on efforts to reduce emissions. (Schäfer *et al*, 2015: 58)

The report understands the moral hazard as potentially occurring at two stages of the SSI process, namely at the decision stage as well as during research and development. Later in the report it states that the relevant agent (the 'who' variable) is society. Of course society is a vague term, as has been indicated in section 4.5, hence it is problematic to use it in this context. Considering the mechanism variable, the report points to phenomena such as increasingly risk-prone behaviour and the creation of dangerous incentives. I understand this as referring to at least psychological failure and ethical failure which can lead to the moral hazard. The above quote from the EUTRACE report makes it clear that the impact variable is to be understood as a negative impact on efforts to reduce emissions. And the badness of this seems to be understood in how it increases the risk of harms from climate change and SSI (Schäfer *et al*, 2015: 59).

The report by the National Academy of Science (NAS) is motivated by the idea that our technical capacity to engage in SRM techniques is greater than our understanding of the ethical, political and environmental impact of such techniques (NAS, 2015: viii). The NAS report understands the moral hazard concern as follows:

> [R]esearch into these proposed techniques could lead to policy-makers deciding to lose focus and/or urgency for reducing emissions. (NAS, 2015: 152)

The relevant stage of SSI for the NAS report is research and development. And the relevant agents are policy-makers. The NAS report does not appear to have an account of the mechanism by which the moral hazard could occur. It is clear that the impact of the hazard is a reduction in mitigation efforts. Yet it is not clear why the report thinks this is bad.

4.7.2 The moral hazard table

The table below summarises how the moral hazard complaint has been formulated by different authors and governance reports

Table 3 The Moral Hazard Table

	Variable					
		Stage	*Who*	*Mechanisms*	*Impact*	*Why*
Author	Corner and Pidgeon	Contemplating geoengineering, and research and development	Individuals and the political level	Psychological mechanisms	Adverse impact on mitigation efforts	Increases the risks of facing harm from climate change
	Morrow	Research and development	Policy-makers	Informational, psychological and ethical mechanisms	Adverse impact on mitigation efforts	Risks a worse state of affairs.
	Lin	Research and development	The public and policy-makers	Psychological mechanisms	Adverse impact on mitigation efforts	Increases the likelihood of facing risks from SSI and climate change
	Bunzl	Research and development, and implementation	—	—	Perverse behaviour objection	—
Report	Royal Society report	Contemplating geoengineering	—	—	Adverse impact on mitigation efforts	—
	NAS report	Research and development	Policy-makers	—	Adverse impact on mitigation efforts	Increases the likelihood of facing risks from SSI and climate change
	EUTRACE report	Research and development	Society	Informational, psychological and ethical mechanisms	Adverse impact on mitigation efforts	Increases the likelihood of facing risks from SSI and climate change

Key: **Red is a point the work is clear on**. **Blue is a point which can be found in the work but which is not explained clearly**

4.7.3 Analysis of the moral hazard table

The fact that the moral hazard is a confused term in the context of SSI is a point which has been made by many authors. Hale observes that 'moral hazard arguments against geoengineering are beset with concerns of ambiguity and vagueness' (Hale, 2012: 129). Table 3 above brings an important – if not the prime – source of confusion into sharp focus. It is simply that none of the authors or reports examined have understood the moral hazard in the same way. It is unsurprising that the term has such vagueness and ambiguity associated with it, given that everyone is using it in a different way. There is a sense in which this is somewhat remarkable, given that these authors are not acting in a vacuum independently of each other and are aware of each other's work; indeed, sometimes they actively criticise each other's work. This should not be taken as a criticism of these authors or reports, but it does indicate that the term 'moral hazard' is still at the stage of being contested.

Not all moral hazard complaints are created equally. Some are much less compelling than others. We can see from the presentation of the literature in Table 3 that one author particularly stands out, namely Bunzl,[132] as does one report, namely the Royal Society report. Both Bunzl's and the Royal Society's accounts are incomplete. They lack an answer to three important variables in terms of who the actors are, the mechanisms involved, and why geoengineering is a bad thing. For this reason, neither Bunzl's nor the Royal Society's understanding of the moral hazard should be adopted due to their incomplete nature.

Table 3 also shows that some authors provide more complete accounts of the moral hazard, namely Morrow, Lin, and Corner and Pidgeon. Each of their accounts appears to be complete due to providing an answer for each of the five variables. This is also true of the account in the EUTRACE report. Despite this, it does not appear that any of them have provided a completely sufficient account of the moral hazard, if it is truly a problem which we wish to address. The reason for this can be seen in the table, in that they all have at least one variable in blue (meaning it is insufficiently explained).

Hence all of these accounts suffer from vagueness on at least one point, which means it would be hard to address any of these accounts of the moral hazard in their current formulation. It is interesting to note how these more complete accounts seem to have emerged at a similar point

[132] Bunzl is not alone in having an incomplete account of the moral hazard. If one was to examine Keith's account (2013), I would argue that it is even less complete than Bunzl's.

in time, whereas the less complete accounts precede them. The stronger accounts were all published between 2013 and 2015 whereas the weaker ones pre-date 2013. This perhaps gives cause to be optimistic about the trend in the hazard literature of at least providing more complete accounts of what the moral hazard complaint is now.

There are some variables on which there appears to be a consensus, namely in regard to what the impact of the moral hazard is, and why it is bad. The impact of the moral hazard is arguably the fundamental starting point of the moral hazard literature. An indicator of this is that even the most incomplete account of the hazard seems to accept what the impact of the moral hazard complaint is. The idea that the fear of a moral hazard relates to it having an adverse impact on mitigation efforts is one which has not been challenged in any of the literature presented in Table 3. Indeed, any prospect of this idea being challenged may have become significantly harder since the work of Hale (2012), who clearly articulated how the impact variable can be formulated.[133]

As for the variable of why the moral hazard is bad, none of the authors presented in the table spends much time articulating why this is the case. Something that this chapter is also guilty of to an extent. It seems clear that many of the articles and reports are motivated by avoiding the risks of climate change and SSI. And whilst it is important to know that there is a moral hazard, there seems to be little of interest to say about why it is a bad thing.[134]

If we consider the variable concerning the stage of SSI at which the moral hazard could occur, there appears to be something close to a consensus that it occurs at the research and development stage. There are two comments to be made about this. Firstly, it is clearly the case that research and development is not the only stage of SSI where the moral hazard could plausibly occur, and Bunzl (2009) at least was aware of this. There is no reason to think that the moral hazard is exclusively a concern of research and development. Indeed, the moral hazard may be even greater at a different stage, such as that of SSI deployment. The dominance of research and development as the relevant stage for this variable could be understood as being a product of where the discourse on geoengineering is currently focused. It appears that we are at the dawn of geoengineering research and development and so it is unsurprising that there is particular interest in the various problems which could arise at the research and development stage, of which the moral hazard could be one. This may be particularly true of the reports

[133] Ben Hale's account was explained in section 4.2.
[134] It should be said that Hale does provide a broad account of why we may think moral hazards are bad by considering consequentialist, deontological and virtue ethics-based concerns with hazards (Hale, 2012: 116-118).

which are being produced with an eye on geoengineering governance. This is why research and development dominates in discussions about the stage of SSI when the moral hazard could occur.

This leaves two variables to consider: those of who the relevant agents are, and the mechanism by which the moral hazard occurs. These are the two variables which appear to be commonly ignored in the early literature on the moral hazard. In the post-2013 literature there also seems to have been a shift in the debate around both of these variables. Much of the post-2013 literature – namely that by Lin, Morrow, and Corner and Pidgeon – appears to be particularly interested in the question of which mechanisms cause the moral hazard. When looking at Table 3 one may think that these authors are in agreement about the psychological mechanisms involved. This is not the case, however. Even though these authors all acknowledge the important role that psychological mechanisms may play, they do not all understand these mechanisms in the same way. On the question of mechanisms, I argued in section 4.4.2 that more information was required from psychologists to generate a clearer account of which psychological mechanisms are in play if the moral hazard occurs.

The final variable is that of who the relevant agents are. As just stated, this variable appears to be ignored in the literature prior to 2013. In relation to the post-2013 literature, 'who' appears to be the perpetually under-theorised variable. Much of the literature and reports point to a political level where the moral hazard could occur, but none of them make it clear how this level is to be understood, and nor do they appear to be sensitive to the dramatic scope for different interpretations of how this level could be understood. This chapter made a move to address this in section 4.3, where it adopted a polycentric account of climate change policy to identify which agents could be relevant to the moral hazard story. I then chose one agent from this account to focus on. This approach is sensitive to the fact that there are many potentially relevant agents and that any account of the moral hazard may need to narrow the scope of possible relevant agents. Given the general level of vagueness about who the relevant agents are and the mechanism by which the moral hazard occurs, it is unsurprising that any satisfactory solutions have yet to be proposed to address the moral hazard in the case of SSI.

The above demonstrates the benefit of the five-variable framing of the moral hazard problem. By categorising moral hazard components in this way, we are able to see many interesting points about the hazard, ranging from where authors tend to disagree to where certain variables are under-formulated. Moreover, we are able to have a sense of whether an author has a clear

understanding of the hazard problem or not. The analysis has provided a way of making sense of why the moral hazard complaint is so confused. Perhaps most importantly, it provides a way of thinking and talking about the moral hazard with clarity.

One limitation of this five-variable analysis is that it fails to draw attention to potential disagreement around any particular variable. For example, two authors could understand the relevant mechanism for the moral hazard as being psychological, but they may have very different accounts of what psychological mechanisms actually are.

In light of this, I shall identify an account of the moral hazard which strikes me as particularly appropriate. The stage of SSI that I will focus on is that of research and development. The reason for this has already been explained, since this is currently where we are in the SSI process and therefore if the moral hazard does exist at this stage, then it is prudent to think about how it could manifest itself and how we may wish to address it. In terms of who, as explained in section 4.3 I shall focus on the formal role-holders of power. And as for mechanisms, I adopt Morrow's account of there being informational, psychological and ethical mechanisms. Regarding the impact and why it is bad, I follow the dominant view in the literature, namely that the moral hazard has an adverse impact on mitigation efforts and is bad because it increases the likelihood of facing a risk from climate change and SSI.

4.8 What role should empirical evidence play?

> [I]f it could be shown empirically that the moral hazard issue was not serious, one of the main ethical objections to geoengineering would be removed. (Shepherd *et al*., 2009: 39)

The quote above from the Royal Society report proposes a condition which, if met, would render the moral hazard objection not serious; namely, if empirical evidence were to show that the moral hazard was not serious. Some academics mistakenly imply that this condition has been met (Bunzl, 2009: 2; Goeschl *et al.*, 2013: 102; Kahan *et al.*, 2015: 203). Reynolds explicitly argues that the empirical evidence undermines the idea that the moral hazard exists, or even shows a reverse hazard (whereby people want to mitigate more due to SSI) (Reynolds, 2015: 175).

In this section I shall argue that such a view is mistaken. This is due to the empirical evidence currently being insufficient to establish whether the moral hazard does or does not exist. This is followed by an argument grounded in the work of Shue on climate change (2010). The argument will state that under certain conditions we have reason to take the moral hazard seriously when faced with such inconclusive evidence about its existence. Consequently, it is argued that if we did not take the hazard seriously due to a lack of empirical evidence, we would be making a moral mistake.

The empirical claim against the moral hazard is often grounded in focus group data which indicate that the general public does not experience a moral hazard (Shepherd et al., 2009: 43; IPSOS Mori, 2010: 2; Mercer *et al.*, 2011: 5; Merk *et al.* 2016). Although there are also more sophisticated experimental methods which have been used to determine whether the public experience a moral hazard. This can be seen in Kahan *et al* (2015), where participants where split into two groups, and only one of those groups was given information about SSI, the other group being the control group. The group who received the information about SSI were in favour of more mitigation action than the control group, indicating the reverse moral hazard. We can grant that the data conclusively establish that the general public does not, never has and will not experience a moral hazard from SSI (note that the claim made that strongly in these articles), but would this be sufficient to say that there is no moral hazard from SSI? Of course not, for the reason provided in section 4.3 of this chapter. Namely, climate policy is a product of a polycentric process in which a variety of different agents are formative of the

policy which is produced. The general public is one of those groups of agents, and probably an important one. But the general public is not the only relevant group.[135] The views of NGOs, businesses, politicians and states also inform mitigation policy.

Recall that I have selected the formal role-holders of power as the agents of interest for my exploration of the moral hazard complaint. As noted, there is a lack of empirical evidence about whether these agents do experience the moral hazard. It also appears that it would be challenging to obtain such data. Not only do the formal role-holders of power tend to change within a ten-year period in democratic countries,[136] but there is also scope for significant variation in the types of people who can hold this office. Therefore, the idea of using empirical evidence to conclusively establish that any potential formal role-holders of power do or do not experience the moral hazard seems very challenging. This would appear to be particularly true if one accepts the cultural cognition thesis as a predictor of whether agents experience the moral hazard, because this would lead to there being very significant variation in the value judgements of all agents who might be or become the formal role-holders of power.

At the moment there is an apparent absence of data about whether the moral hazard would be experienced by all the relevant agents – or even just the agents which this thesis is interested in. Hence it appears that we cannot yet make any empirical claims about whether the moral hazard will or will not occur. If such data do exist in the future, then one could reconsider how seriously the moral hazard should be taken in light of this empirical evidence.[137] In the meantime, however, there is the question of how seriously the moral hazard should be taken, given the inconclusive nature of the empirical evidence available. In the following sub-section I shall argue that the moral hazard should be taken seriously despite the absence of empirical evidence about whether the formal role-holders of power experience the hazard or not.

---

[135] The policy level is subject to pressures which individuals are not. For example carbon intensive industries may well apply pressure to the formal role holders to get them to succumb to the logic of the hazard. Moreover formal role holders of power are subject to complicated polarising policy dynamics which individuals are not (Boven *et al*, 2018), whilst this thesis does not explore these dynamics, these dynamics clearly put different types of pressure on the formal role holders of power which mean it is plausible for them to reach different conclusion than that of the general public on policy issues, and this includes their views on SSI.

[136] That is to say there will new members of the legislature and the executive.

[137] Although the pursuit of such data may be misguided, for it is doubtful whether it could establish the absence of the moral hazard conclusively. Some reasons as to why this is the case are provided in the previous footnote.

4.8.1 Why should the moral hazard be taken seriously if we are uncertain whether it will happen?

To make the case for the moral hazard being taken seriously in the absence of conclusive empirical data about its existence, this sub-section will draw upon the work of Henry Shue. Shue provides an account of why action to mitigate the harms of climate change should happen even if we do not know the precise probability of the harms which could occur from climate change (Shue, 2010). In short, he argues that action to address climate change should happen if (1) there are massive losses at stake; (2) the likelihood of the loss is significant; and (3) the cost of acting is non-excessive. In this sub-section I shall make two arguments. Firstly, that Shue's account is a compelling one, and secondly, that the case of SSI and the moral hazard does meet Shue's criteria for action. By doing so this sub-section will explain why we should take the moral hazard seriously despite the inconclusive nature of the empirical data we have about it.

One of the great appeals of Shue's account is that it is action guiding. Shue asks us to consider three conditions, and if each of these conditions is met, then there is an imperative to address the problem. This is much clearer than alternative accounts of why action should happen in the face of climate change, such as generalised accounts of the precautionary principle.[138] Additionally, there is also a precedent of using Shue's framework in the context of geoengineering, as presented by Hartzell-Nichols (2012: 160) and McKinnon (2018: 446).

The account of the badness of the moral hazard which was considered in section 4.4 shows how the hazard meets Shue's first condition of massive losses being at stake. In short, the moral hazard has increased the likelihood of some of the worst harms of climate change and SSI happening. McKinnon vividly makes the case for this condition being met by focusing on the risk of the termination shock posed by SSI (McKinnon, 2018: 448). This is one among many of the sources of badness for the hazard which was pointed to in section 4.4. If the termination shock is sufficient to meet Shue's first criterion, then the reasons provided throughout this chapter likewise easily meet it, for they can be understood as the termination shock plus many other risks besides.

[138] This is not an argument against precautionary approaches. Shue's account is also a formulation of the precautionary principle, yet it is quite specific and clear.

Shue's second condition is the most contentious one, that of the likelihood of the loss being significant. For Shue, this condition is met on the basis of two components. The first component is that the mechanism by which these losses can occur needs to be well understood (Shue, 2010: 148). The second component is that the conditions under which this mechanism would function are accumulating (Shue, 2010: 48). The first component was introduced in section 4.6 where the three causal mechanisms of informational failure, cognitive failure and ethical failure were explored. The following paragraphs will now consider how much these mechanisms are understood and whether they are well understood. I will argue that the conditions under which the psychological mechanisms would operate are not currently well enough understood to meet Shue's criterion but that the informational and ethical mechanisms are.

The relevance of the psychological mechanisms is clear, given the role that such mechanisms play in people making judgements about risk (Kahan, 2015). Nonetheless, the process of the psychological mechanisms is not yet fully understood, although the work of Lin provides reason to think that the overconfidence and optimism biases will be relevant. However, as explained in section 4.4, further research is required if we are going to be confident that these and other mechanism do have an impact on the moral hazard. In order to judge whether the conditions under which these psychological mechanisms would function are accumulating, one would probably need to be more certain about their content and how exactly they operate. Therefore, given our current level of knowledge, it does not appear that the psychological mechanisms meet Shue's criterion because we do not seem to have knowledge of the conditions under which they would function.

As for informational failure, two sources were identified, those of ignorance and false belief. The presence of ignorance about the risks, efficacy and promises of SSI was rationalised as the default state of affairs for the formal role-holders of power, given the specific and technical nature of this type of knowledge. False belief was not explored in detail. Ignorance is considered to be an important aspect of the informational failure mechanism, although further research could elaborate on how this mechanism functions in the case of SSI. There is already a well-established literature in economics on the role of ignorance as a causal mechanism for the moral hazard to occur. Nonetheless, this thesis does not provide a completely well-understood account of informational failure due to the lack of understanding about the operation of false belief.

Despite this, we can consider the second component of Shue's second criterion by asking whether the conditions leading to ignorance are accumulating. It is odd to ask whether the conditions for ignorance are accumulating, given that ignorance about the risks, efficacy and promises of SSI is taken to be the default state of affairs for many formal role-holders of power. Even if the conditions for ignorance are not accumulating, they appear to be sufficiently present that there is certainly a danger of a critical number of formal role-holders of power being ignorant about the risks, efficacy and promises of SSI.

The final mechanism to consider is that of ethical failure. In section 4.6, moral corruption was identified as a mechanism by which ethical failure can occur. The conditions for moral corruption are accumulating due to the presence of the perfect moral storm (Gardiner, 2011). There is little reason to think that the elements of the perfect moral storm have reduced in strength. There are still significant disparities between the global rich and global poor which are constitutive of the global storm. There is still a strong divergence of interests between present and future generations which is a key feature of the intergenerational storm. And our best theoretical tools from economics, politics and philosophy seem to struggle to make sense of the problems which are posed by climate change and SSI, such as what is owed to non-human nature. Given that the perfect moral storm still seems to be in play, and that moral corruption is compounded by the storm, and that ethical failure in this case can be a product of moral corruption, it seems reasonable to think that the conditions for ethical failure to occur are accumulating.

So has Shue's second condition been met? The answer is 'yes', in part. The mechanisms of informational failure and ethical failure do appear to tick both of Shue's boxes; they are well understood and the conditions for their functioning seem to be accumulating. The same cannot be said of the psychological mechanism, however, since the content of this mechanism is not yet well understood in the context of SSI. Consequently, we cannot judge whether the conditions for this mechanism are accumulating.

Shue's third condition is that the cost of addressing the problem is non-excessive. Whether a cost is excessive is determined by comparing it to the magnitude of the potential loss and the other important demands on the resources which could be used to pay for addressing this loss (Shue, 2010: 149). As of yet, the question of excessive cost is not one which this thesis can answer, for it has yet to consider the possible strategies for addressing the moral hazard. That

is the task of the subsequent two chapters. Of the two solutions considered in this thesis, it will be argued that one does impose an excessive cost, yet the other does not.

To summarise this section, there is an important role to be played by empirical evidence, yet so far it appears to have been used poorly: by Reynolds (2015), for example. It has been used poorly by focusing on a specific group of agents such as the general public and then generalising from this to say that the moral hazard does not appear to exist. Additionally, the type of evidence which is required has not been correctly conceptualised. It is not simply a question of establishing the probability of an agent experiencing the moral hazard. Instead, it is more appropriate to adopt Shue's approach and seek evidence to explain the mechanisms by which the moral hazard may occur, and also to investigate whether the conditions for these mechanisms to operate are accumulating.

There is the further question of what should be done in the absence of conclusive evidence that these conditions are being met. The choice appears to be: (1) do nothing to address the moral hazard until such evidence exists; or (2) act in the knowledge that such action may not be required if the evidence ultimately shows that Shue's second condition of the loss being significant is not met. Given Shue's third condition that action should be taken if the costs are non-excessive, this thesis adopts the approach that it is at least worth understanding how we could respond to the moral hazard in case it does exist, as long as the ways to address the hazard do not fall foul of the third condition.[139]

[139] This approach appears to be consistent with the reports on geoengineering governance, which flag the moral hazard as a problem which may need to be addressed.

4.9 Conclusion

This chapter has sought to provide a degree of conceptual clarity as to what the moral hazard concern is, which has thus far been absent from the philosophical literature and governance reports on SSI and the moral hazard. To do this the chapter engaged with the keys works on the moral hazard problem in the context of SSI. This engagement has allowed the thesis to identify five components of a complete moral hazard complaint, namely: the stage of SSI, who the relevant agents are, the mechanism by which the hazard occurs, what the impact of the hazard is, and why the hazard is bad.

The five-variable understanding of the moral hazard complaint has allowed the chapter to achieve its goal of providing conceptual clarity as to how the hazard should be understood. This understanding of the moral hazard offers clarity for three reasons. Firstly, it makes it possible to identify incompleteness in a particular account of the moral hazard by noting the absence of one or other variable. Secondly, it draws attention to any vagueness in an account of the moral hazard. Such vagueness seems to be particularly common in the literature when it comes to the 'who' variable. Thirdly, it highlights divergent understandings of the moral hazard complaint when accounts of the hazard fill out the variables in distinctly different ways. Interestingly, the governance reports on SSI seem to suffer from all three of these shortcomings.

This chapter proceeded by presenting an account of SSI which overcomes the first two of these concerns. That is to say, it is a complete account and it is not vague. According to this account, the moral hazard occurs at the research and development stage of SSI, the relevant agents are the formal role-holders of power, the casual mechanisms include informational failure, psychological failure and ethical failure, the impact of the hazard is that it has an adverse effect on mitigation efforts, and this is bad due to the fact that it would increase the likelihood of risks from climate change and SSI.

The chapter also performed the further task of considering what empirical evidence would be required for the moral hazard concern to be taken seriously. In doing so it expressed scepticism about how empirical evidence has been used by people like Reynolds (2015) to cast doubt on the existence of the hazard, this was done by observing that the evidence is only about the general public, not any other relevant agents such as the forma role holders of power. It was argued that Shue provides a more plausible account of what empirical evidence would be

relevant, and that there is a need for such evidence to be sought. Despite the absence of some of the necessary empirical evidence, this chapter reasoned that it is still worth taking certain steps to address the moral hazard (such as understanding how it can be addressed) insofar as such moves are compatible with Shue's third condition of non-excessive cost.

5 Secrecy and the Moral Hazard: Effective but not Desirable

5.1 Introduction

In the previous chapter I considered how the moral hazard complaint can be understood. In this chapter I build upon this understanding of the hazard to explore the possibility of secrecy as a tool to address the moral hazard which is presented by SSI. To do this I will explore an insightful account of secrecy which is provided by David Pozen. In order to understand what Pozen's account of secrecy shows, I will apply it to the case of the development of the first atomic bomb, known as the Manhattan Project.

Consideration of the Manhattan Project is particularly useful here due to the level of knowledge that we have about how secrecy in the Manhattan Project functioned. There are also some interesting similarities between the case of SSI and the Manhattan Project in terms of the development of a novel technology which is potentially harmful and requires extensive expertise. This will show how secrecy has functioned in practice. I will use Pozen's framework to consider different scenarios of secrecy for SSI and what secrecy might look like in different contexts, before considering whether secrecy would address the moral hazard problem. I will argue that in a scenario in which secrecy does address the moral hazard, we have a plethora of powerful reasons to be sceptical about it. These reasons include considerations of publicity, procedural justice moral corruption and co-benefits. Therefore I will conclude that secrecy is not a morally desirable tool to address the moral hazard problem.

Secrecy is remarkably powerful; it can be used to inflict harm as well as to protect individuals from significant harm. In the case of the moral hazard, there is a *prima facie* reason to think that secrecy may well address the moral hazard concern. As was established in Chapter 4, the moral hazard concern of interest to this thesis is that the research and development of SSI may have an adverse impact on the climate change mitigation efforts undertaken by the formal role-holders of power.[140] For this adverse impact on mitigation to occur, the formal role-holders of power require knowledge about the research and development stages of SSI in order for the casual mechanism to be activated. Therefore, if the formal role-holders of power lack the

[140] It may seem odd to think of secrecy as a tool to be used against formal role-holders of power given that they are often the keepers of secrets, nonetheless I think this the oddness also makes it an interesting case to consider, particularly given its potential to work as a way to address the hazard which this chapter will demonstrate.

relevant knowledge to form an expectation about SSI, then their behaviour will not change.[141] Hence we have a *prima facie* case that secrecy about SSI would prevent a moral hazard occurring.

This power of secrecy has not been noted by many other scholars. An exception is Albert Lin, who acknowledges that:

> [L]imiting the visibility of geoengineering efforts may offer one mechanism for countering people's tendency to compensate for risk. Such a strategy would be troubling and should be avoided, however, as it is contrary to fundamental democratic values of transparency and public deliberation. (Lin, 2013: 709)

This seems to be the extent of the literature on secrecy and the moral hazard in the context of geoengineering. This dismissal of secrecy in the literature on SSI and the moral hazard leaves an unexplored area which this chapter aims to fill. By expressing the logic in favour of secrecy, we will have a clearer idea of what Lin is rejecting on the grounds of fundamental democratic values.[142] More importantly, we will have an image of how secrecy could be used to address the moral hazard concern. The chapter will then provide four reasons for finding this degree of secrecy undesirable due to: being an obstacle to procedural justice, enabling corruption, enabling moral corruption, and limiting the opportunities for co-benefits.

Secrecy may appear to be at odds with chapter 3 of this thesis, where the case in favour of transparency, publicity and accountability was made in order to address moral corruption. It is true that these ideas are not compatible with secrecy (a point explored in more detail towards the end of this chapter), yet this is only a problem if one thinks that moral corruption is sufficiently serious. This thesis does not weigh these problems against each other. Instead it seeks to understand how these respective problems can be addressed. The tension only arises if one wishes to address both of these issues, although the thesis has sought to provide reasons why both these problems are worth taking seriously.[143]

---

[141] Of course, an agent's behaviour can still change for reasons which are not related to their expectations about SSI.

[142] This can be understood as the liberal theory objection to secrecy which is explored by the likes of Thompson (1999).

[143] Whilst I hold the view that it is worth addressing both of these problems, I cannot assume this would be true for everyone who seeks to address the moral hazard problem (we may even think that this is unlikely due to the fact that the moral hazard issue seems to be better known than the problem of moral corruption). Moreover one may have an alternative way to address the problem of moral corruption which does not rely on transparency, publicity and accountability. This alternative solution may well be compatible with secrecy. This is another reason not to be concerned by the apparent tension between the moral corruption chapter and the current chapter.

5.2 What is secrecy?

To help conceptualise secrecy, I will engage with the work of David Pozen (2010).[144] The importance of this section is that it provides a degree of conceptual clarity about what secrecy is. This is necessary for the chapter to be able to consider what secrecy could look like in the case of SSI governance and whether it is desirable or not. Pozen provides a sophisticated exploration of 'deep secrecy', a term which he traces to Kim Lane Scheppele (Pozen, 2010: 262). According to Scheppele, deep secrecy exists when the secret is unimaginable to the individual from whom it is being kept secret. An example of this would be a well-planned surprise party (Pozen, 2010: 263). This is in contrast to shallow secrecy, which relates to secrets one might suspect to be the case. So if you expect your friends to have organised a surprise party for you because they do so every year, then the secret is a shallow one.

The term unimaginable appears to be misleading and mistaken in this context. Events such as surprise parties are imaginable. The secret is not dependent upon whether the party *can* be imagined, it is dependent upon whether something *is* imagined and expected. Therefore it is more appropriate to understand a deep secret as being something which is not imagined, as opposed to something which is unimaginable. I will use this modified understanding of deep secrecy in this chapter.

Notably, Pozen observes that the idea of deep secrecy is almost ignored in the governance literature on transparency (Pozen, 2010: 262). This is because the literature on deep secrecy follows Scheppele, whose level of concern with deep secrecy was transactions between individuals rather than at the level of the state. An exception to this is found in the work of Amy Gutmann and Dennis Thompson (1996), who apply the concept of deep secrecy to the state level Gutmann and Thompson argue that a shallow secret at the state level is when the citizens know that the state level secret exists but do not know the content of the secret. In contrast, they understand a deep secret at the state level as one where the fact that there is even a secret is concealed from the citizens. An example of a shallow secret at the state level is the codes to activate nuclear weapons; the citizens know that the secret exists but the code itself is still a secret.

[144] There are several good works which engage with secrecy such as that by Bok (1982), De Lazari-Radek and Singer (2010, 2014), and Sagar (2013). The reason for engaging with Pozen here is due to his lucid focus on secrecy in the context of governance.

The purpose of introducing the Gutmaan and Thompson account of deep secrecy was to draw attention to how Pozen builds on them, Whilst Pozen holds the same broad understanding of what deep secrecy is he spends much more time conceptualising secrecy. Pozen, in contrast, is much more concerned with how to conceptualise secrecy (Pozen, 2010: 275). Consequently, Pozen's article is of greater use in this part of the chapter, given that we are concerned with trying to conceptualise what secrecy is. In order to understand the depth of a secret at the state level (i.e. whether it is a deep secret or a shallow one), Pozen provides a set of challenges which need to be considered in relation to the practice of keeping secrets at the state level (Pozen, 2010: 266). I shall now explain Pozen's framework and then apply it to SSI governance.

Pozen argues that current applications of the idea of shallow and deep secrecy by the likes of Scheppele are unsatisfactory for three reasons (Pozen, 2010: 265-267). If we examine Pozen's three reasons for concern with the way the idea of secrecy has been applied to date, we can be sensitive to these considerations when building up a clear account of secrecy. The first of Pozen's concerns is that current understandings of secrecy are not sensitive to the fact that the depth of a secret exists on a continuum, as opposed to being a binary concept. To illustrate this, Pozen points to how one can have a vague sense that a secret exists, and how that sense can get ever stronger, and as this happens the depth of the secret is reduced, yet we may still classify it as a deep secret.

Pozen's second concern about how secrecy is classically applied is that thinkers like Scheppele consider secrecy as existing only between two agents: the keeper of the secret and the subject from whom the secret is kept. Pozen observes that this is clearly not true at the state level, where there can be multiple secret-keepers and multiple subjects from whom the secret is kept (Pozen, 2010: 266). This point will be seen more clearly later in this chapter when secrecy is applied to the case of SSI governance, given the number of actors who could partake in the secret-keeping and the number of subjects from whom the secret could be kept.

Pozen's third concern about secrecy is what he calls the 'problem of specification' (Pozen, 2010: 267), which needs to be addressed in order to understand whether something is actually considered secret or not. Pozen's point is that whether X is a secret or not depends on how it is understood, and he gives the example of the Manhattan Project to illustrate this. Basically, if one were to adopt a broad standard for the Manhattan Project, such as the US developing more powerful weapons during World War II, then it is unlikely to be considered a secret, since many people would have imagined that the US would undertake the development of more

powerful weapons when it was at war. Yet it is much more doubtful that people would have imagined the development of the nuclear bomb. Therefore, if one specifies that an account of the Manhattan Project secret should include detailed knowledge of the nuclear weapons which were being developed, then the Manhattan Project was indeed a very successful secret. Hence, depending on how much about the Manhattan Project was deemed a secret, the secret-keepers were either very unsuccessful or extremely successful at concealing the secret. Pozen's point is that to judge the success of any secret, one requires an account of what exactly that secret is.

When considering a case of secrecy, Pozen proposes that the secret should be evaluated against four indices. By considering these indices, he thinks that we will have a picture of secrecy which is sensitive to the three concerns about secrecy just explained. He sets out the four indices as follows:

> The most productive way to conceptualize the depth of a state secret, I propose, is to evaluate it along four main indices, reflecting (1) how many people know of the secret, (2) what sorts of people know, (3) how much they know, and (4) when they know. (Pozen, 2010: 267)

Unfortunately, Pozen fails to offer a justification for the relevance of these indices; he merely asserts that they are the most productive way to conceptualise deep secrecy. Nonetheless, despite the lack of argumentation for his indices, I think they are promising for our purpose here. I shall now outline the relationship between these indices and Pozen's three questions, and by doing so offer a justification for using Pozen's indices as a way of understanding the depth of a secret.

The first two of these indices are clearly meant to account for Pozen's second concern about there being multiple agents, since the answer to these indices will explicitly highlight the number and different groups of agents who know of the secret. The third and fourth questions are designed to obtain a clear picture about how much these agents know and when. This appears to speak to Pozen's first concern about understanding secrecy on a continuum; by identifying how much an agent knows, it is possible, in Pozen's terms, to locate where on the secrecy continuum they are. This also seems to have a bearing on Pozen's third concern about the problem of specification. If one is trying to determine the specification of a secret, then knowledge about where the secret-holders are on the secrecy continuum would allow us to specify the depth of the secret in terms of which agents are or are not included in knowing about it.

I shall now clarify the location of secrecy, the previous chapter expressed an understanding of the moral hazard problem as occurring at the stage of research and development. The question then naturally becomes what would need to be secret for the research and development of SSI to be a secret? The obvious answer is for the research and development itself to be a secret. Here I hold a slight more complicated view, whereby I think it is also the governance of SSI which needs to be a secret for the research and development of SSI. The reason for holding this view is based on the thought that if the relevant agents were aware of SSI governance then it is easy for them to imagine and expected SSI to be research and development, so SSI at most could only every be a very shallow secret form them. Yet if they are unaware of SSI governance it requires more imagination for them to think that SSI is being researched and developed, therefore the secret is deeper to these agents.

I shall now comment on each of the indices identified by Pozen as it applies in the case of SSI. Once I have done so the relevance of these indices for Pozen's three concerns in relation to SSI ought to be clearer.

5.3 Answering Pozen's questions

5.3.1 How many people know about SSI governance?

'Three may keep a secret, if two of them are dead.' Benjamin Franklin

Pozen starts with the simple observation that the number of people who know a secret will have an effect on whether the secret is compromised or not (Pozen, 2010: 268). In other words, if fewer parties are aware of a secret then the total number of opportunities for the secret to be revealed, leaked or hinted at is reduced. The implications for SSI governance are clear, namely, that as more people become aware of the secret of SSI, the greater the risk of the secret being exposed.[145] Therefore, as more people know about it, the shallower the secret becomes.[146]

5.3.2 What sorts of people know?

Pozen provides a particularly helpful point when considering what sorts of people know about the secret, which is that governments are not monolithic (Pozen, 2010: 269). This points to the idea that within a state there will be some members of the government who are aware of the secret of SSI governance, and some who are not. Who is and is not aware of particular secrets in a state will in part be a product of how secrecy operates within that state.[147] This variation could exist across institutions, in terms of members of the executive, legislature, judiciary, civil service, military, universities and private research groups.

5.3.3 How much do they know?

Following on from Pozen's previous question, once we have identified what sorts of people know *of* the secret, there is the question of *how much* they know. The point is a simple yet

[145] There are different ways of understanding what the secret of SSI governance is, as will be made clear. This is simply a general comment about how it is easier for a secret to be revealed if more people know it, and that this is true independent of the content of that secret.

[146] This assumes that everyone is equally likely to keep the secret, which may or may not be true; it may well be the case that a larger group of trustworthy people are more likely to keep the secret than a smaller group of untrustworthy people.

[147] The secret in this case is understood in terms of the agents having a big or a complete picture of what is happening.

important one, which is that different agents can know of a secret to different degrees. To think otherwise is to make the mistake of thinking of secrecy in binary terms as opposed to a continuum. [148] When we conceive of secrecy as a continuum it makes sense that people can be aware of a secret to different degrees, hence there can be variation in how much people know. For example, consider one variable in the case of SSI governance: some agents may be aware that a governance institution for SSI exists, others may have knowledge about the functioning of SSI governance, and others still may know every decision that the governance regime makes.

5.3.4 When do they come to know?

Pozen explains that a secret may be deep when it is initially conceived, yet it may become ever shallower over time (Pozen, 2010: 27). An example in the case of SSI could be around deployment. If SSI were to be deployed, then the institution which governs SSI would probably not expect the deployment to be a secret. This is because the use of SSI is detectible. Some groups may be able to detect the deployment of SSI almost immediately, whilst others may take months or years to do so, but the secret will become shallower the further we get from the point in time at which SSI is deployed. This is due to SSI being easier to detect the further it is from the time of initial deployment.[149]

5.3.5 Pozen's understanding of secrecy

Pozen's four questions provide a starting point for understanding the depth of a secret which is held by a state. These questions exist in response to the three challenges that Pozen identified in relation to understanding the depth of a secret. These challenges are: (1) the fact that secrecy lies on a continuum; (2) having multiple agents who do or do not know the secret; and (3) the problem of specification, which is the question of how the secret is understood. The use of Pozen's framework was defended towards the end of section 5.2 above. As explained, my

[148] Pozen does not provide any criteria to identify when a deep secret becomes shallow. Instead, he provides these questions as a way to think about each case of secrecy. By answering these questions we should be well placed to understand where a case of secrecy lies on the secrecy continuum.
[149] For the time taken to detect the use of SSI, see (Lo *et al,* 2016; Lo et *al,* 2018; MacMartin, 2019).

purpose is not to challenge Pozen's framework for understanding secrecy, but rather to use it as a starting point to make sense of secrecy in the case of SSI governance.

5.4 Applying Pozen's four questions to the case of the Manhattan Project: understanding secrecy

In this section I will apply Pozen's four questions to the case of the Manhattan Project. By showing how secrecy worked within this project we can gain insights into how it might work in the governance of SSI. Prior to this, I shall justify the relevance of the Manhattan Project for thinking about secrecy in the case of SSI governance.

The relevance of this case is twofold. Firstly, it is a case of secrecy where there is a lot of information available. Importantly, there is enough information to provide sufficient answers to each of Pozen's four questions. This means that the Manhattan Project meets some basic criteria in regard to being applicable in this case. Secondly, the Manhattan Project faced certain challenges similar to those which would occur in relation to the governance of SSI. SSI requires the development of a new technology which is potentially harmful and requires specialist expertise, just as the Manhattan Project did. This will make some features of the project particularly relevant for the case of SSI, such as compartmentalisation (which will be explored in this section). There are of course significant differences between these two cases as well. Not least, that SSI is not a weapon.[150] Another note worth difference is how different society is now in contrast to the 1941 world in which the Manhattan project began. For example the nature of communication is dramatically different (consider all the communication done via the internet), the different nature of communication creates challenges for secret keeping which simply did not exist when the Manhattan project

My claim here is not that a Manhattan project equivalent is possible in the case of SSI, which would be foolish; the claim is rather that it is sufficiently similar that we can learn about the practice of secrecy. This subsection concludes by drawing attention to these lessons, such as those of compartmentalisation, strong leaders and the capacity to generate funds without revealing what they are being spent on. The information we have about the Manhattan Project and these certain similarities are the reason why I choose to use the Manhattan Project as my case of secrecy, although other projects are surely also suitable. For example, one could run the same analysis with the case of the code-breakers at Bletchley Park.[151]

[150] Although SSI may be used for military purposes (Cairns and Nightingale, 2014).
[151] This was also done in an earlier draft of this chapter.

The Manhattan Project holds a unique place in history: it was the first project to successfully develop the atomic bomb. The project began in the United Sates at the end of 1941 and it was led by General Groves. This means that it was led by the military rather than scientists, although the project had brilliant scientists working for it such as Enrico Fermi, Richard Feynman and Robert Oppenheimer.[152] In May 1945 the bomb was successfully tested, and it was then used in August 1945. The scale of the project was huge, costing nearly $2 billion (approximately $30 billion in today's terms), and it employed over 100,000 people (*Jones, 1985).*

5.4.1 How many people knew?

I understand this question in terms of what groups of agents knew about the Manhattan Project, as opposed to how many individual people knew. The reason for this is that it would be unhelpful and difficult to calculate the precise number of people, to the point of it being unknowable. In contrast, we can identify which groups knew the secret, and knowledge of this is helpful for understanding the nature of the project's secrecy. For the purposes of this analysis I identify three relevant groups: the general public, employees of the project, and the government. There are different ways of dividing agents into groups, but the claim here is that the division into these three groups is appropriate for understanding secrecy in this case, given that the purpose is to provide tools for thinking about secrecy in relation to SSI governance.[153] This will also become clear by the understanding of secrecy which emerges in this section when we focus on these groups.

5.4.2 What sorts of people knew?

This question invites us to consider the composition of the groups identified by the previous question. The question here is one of whether there is any relevant variation between these groups in terms of the roles that they perform. With regard to the general public, there do not

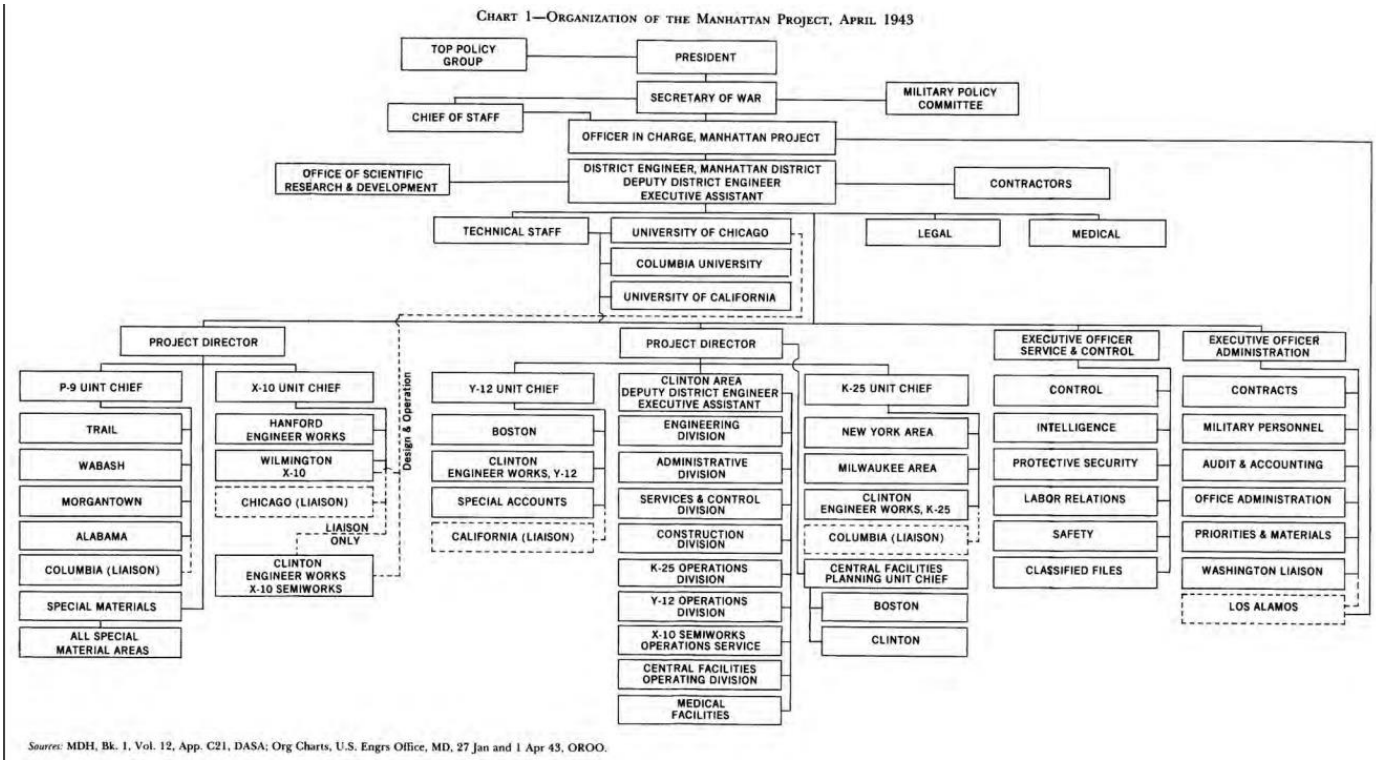
[152] Eight members of the project received a Nobel Prize at the time it started and a further 12 would receive one after the project was completed (Norris, 2004: 66).
[153] This is the reason why foreign powers are not included in this analysis, because although they are relevant to understanding the Manhattan Project generally, they do not play a significant role in the case of SSI governance at the global level.

appear to be any relevant variations in the composition of this group, whose role is that of citizens being protected by the state.

When considering what sorts of people employed by the project knew the secret, the question invites us to consider what different roles existed within the project. If we consider all those who were employed by the Manhattan Project, there is a huge amount of relevant variation. At the top of the project's hierarchy there were General Groves and leading scientists such as Oppenheimer. Lower down the organisation there were people who were directing smaller projects as part of the whole, and then there were other employees on these projects, since as with any kind of organisation there were administrative staff as well. Groves performed the role of leading the project, whilst lower down the organisational tree there were many scientists and universities carrying out separate tasks, such as the University of California which helped Ernest Lawrence to develop the electromagnetic process for separating atoms (*Britannica Academic*, 2019).

It is clear that a project which employs 100,000 people is likely to have great variation with regard to the sorts of people who are part of the project. This is illustrated by Figure 4 below, which sets out the organisational chart for the Manhattan Project (Jones, 1985: 88–89).

Figure 4: The Manhattan Project Organisational Chart



When considering which actors within the government knew about this project, the organisational chart is also helpful. It shows that some important members of the government were aware of the Manhattan Project. These were the Secretary of State for War, the corresponding Military Policy Committee, President Franklin D. Roosevelt (FDR), and his top policy group. This is a relatively small number of actors from the government side, when we consider that the government of the United States includes the two legislative chambers of Congress (the House of Representatives and the Senate), which do not appear to have known about this project. In other words, the vast majority of people in government were ignorant of the project, with some noteworthy expectations who appear to be individuals trusted by President Roosevelt.

5.4.3 How much did they know?

There was clearly a wide range of employees involved in the Manhattan Project. Indeed, this range of employees poses one of the biggest and most intriguing mysteries about the project: how does a project manage to employ 100,000 people and still maintain secrecy? A key component of achieving secrecy in the Manhattan Project was the strategy of compartmentalisation. Compartmentalisation is the division of a task or project into particular pieces. For the Manhattan Project this meant keeping different teams with different roles separate from each other (Vermeir and Margocsy, 2012: 159).

For example, the vast majority of physicists and engineers on the project did not interact. Part of the logic for compartmentalisation in the case of the Manhattan Project was that General Groves did not want any individual to know too much about how the whole project worked. Little more than a dozen people were really aware of what the project in its entirety was doing (Jones, 1985: 89). Therefore, the answer to our question is that most people knew very little. Many people working in teams may have had a sense that they were working on weapons research and development, but they did not know about the nature of the weapons that they were contributing to, and this was very much due to compartmentalisation.

Other techniques were used to limit knowledge about the progress being made on atomic energy by the Manhattan Project in regard to people both within and outside the project. For example, American scientific journals agreed to stop publishing articles on atomic energy (Jones, 1985: 26). This resulted in the American atomic energy community effectively going silent, so that those outside the community no longer knew about their advances in atomic energy. This effectively restricted knowledge of the advances in atomic energy to those who were sufficiently high up the organisational tree of the Manhattan Project (Jones, 1985: 26).

The effect of this would be that those lower down the tree, or who were not part of the organisational tree at all, would not be informed of these advances. Therefore, it appears that this decision contributed to keeping significant knowledge about atomic energy secret from many within and external to the project.[154] This culture of secrecy continued on into the Cold War. It resulted in classified journals existing, where scientists could peer review and access each other's work only if they had a sufficient level of clearance.

[154] Interestingly, the Manhattan Project also provides some lessons about detecting secrecy. Georgii Flerov (a Soviet nuclear physicist) observed the silence from the American atomic energy community and predicted that this meant they were developing the atomic bomb. So if parts of the SSI community (small as it is) were to go silent, this could be an indicator that they are now performing research in secret. That is a possible answer to an important question, namely: what would indicate the existence of a secret SSI community?

Regarding variation in knowledge at the government level, clear variation has been identified. There was a small political community, apparently led by the President, which was aware of the Manhattan Project, in that they knew what was being developed. Whilst a few other political actors did too, Congress certainly did not know much about the project. President Roosevelt got Congress to provide the $2 billion dollars of funding without revealing the details of the programme. The extent of knowledge within Congress seems to be that it was providing $2 billion for a weapons research and development programme which would be helpful to the war effort. This implies that the difference between the knowledge that Congress had and that of the general public who suspected that such programmes existed is that Congress knew the cost of these programmes.

5.4.4 When did other groups come to know?

In the context of the Manhattan Project, Pozen's final question asks when other groups learnt about the secret. In one sense the secret of the Manhattan Project was revealed in August 1945 when the atomic bomb was first used, since at that point the existence of the bomb was no longer a secret. Yet there were plenty of secrets connected to the Manhattan Project which were kept secret after 1945. Knowledge about how the bomb functioned was still a closely guarded secret because the US did not want other countries to develop this technology. Over time, however, more and more information has been revealed to the public, to the extent that we know about the Manhattan Project and we know how the project was organised and what was learnt on the project. In fact, the official history of the Manhattan Project has now been released, which provides a huge amount of information about it: 36 volumes in all.[155] This illustrates two points: firstly, the secret may not last forever, and secondly, what is secret may change over time as people's access to certain knowledge changes.

5.4.5 Some lessons from the Manhattan Project about secrecy as a governance practice for SSI

[155] These volumes can be found at https://www.osti.gov/opennet/manhattan_resources [accessed in January, 2018]

This section has sought to address Pozen's four indices of an account of secrecy using the example of the Manhattan Project. The purpose of Pozen's four questions is to build a picture of secrecy which is sensitive to the shortcomings of how secrecy has been understood in the past. As explained in section 5.2, these shortcomings concern whether a secret is understood in binary terms as being simply deep or shallow, being sensitive to the plurality of secret-keepers and subjects who are kept ignorant of the secret, and the problem of specification. Pozen's indices fulfilled their purpose, for the picture of secrecy which they provide in relation to the Manhattan Project does not seem to suffer from these shortcomings. This can also be seen in the following paragraphs, where I shall draw attention to some of the empirical points from this analysis which will be of use for understanding how secrecy can function in the case of SSI governance.

The Manhattan Project relied on extensive compartmentalisation, and the depth of its secrecy is in part a product of the function of compartmentalisation.[156] Compartmentalisation is a strategy which both enabled the Manhattan Project to employ a great number of people and limited the risk that any one individual would be able to expose the secret of the project in its entirety. A noteworthy feature of compartmentalisation is that it seems to require a few visionaries who can guide the project whilst everyone else is kept in the dark about the true nature of the project. In the Manhattan Project this was General Groves and his trusted scientists. A further relevant lesson relates to not publishing scientific advances in public access journals.

The ability of FDR to generate the funding for the Manhattan Project without telling Congress about it in detail shows that in some contexts funding can be obtained for these types of projects without the secret of the project being revealed. He was able to do this due to the presidential power to allocate military funds, as permitted by the Military Appropriations Act of 1944 (Jones, 1985: 116). This shows that at least some states have mechanisms for allocating funds in particular circumstances in such a way that their legislature is unaware of how these funds are being spent, despite having the power to approve the national budget. If one wants the governance of SSI to be a secret, there may be a comparable question of how such governance can be funded whilst limiting types of knowledge about SSI from those who provide the funds.[157]

[156] This strategy is not unique to the Manhattan Project; it was also practised in Bletchley Park (Grey, 2014: 112).
[157] The type of knowledge you would want to limit is contingent upon the account of SSI secrecy that you aspire to. This will be seen in the coming section.

It is interesting to note that identifying how much the government knew is one of the trickiest tasks from this section. What is clear is that FDR was aware of the Manhattan Project, and General Groves who led the project was directly reporting to him and was meant to be keeping FDR well-informed about the project. Yet knowledge of this project within the wider government seems to have been very limited. This is indicated by Figure 4 above showing the organisational chart of the Manhattan Project, where there are very few government officials with a role within the project. Comparable arrangements may be possible in the case of SSI. Which is to say that the head of government may know about how a project is developing without having a detailed idea of what actually happens within the project.

Public information about the prospects of nuclear weapons seems to have been non-existent prior to and during the Manhattan Project. The press did not run stories about the possibility of nuclear weapons being developed (Jones, 1985). The lack of public information on nuclear weapons helped to keep the Manhattan Project a secret from those external to the project. One of the great challenges of a Manhattan Project-esque level of secrecy for SSI is the amount of publicly available information which already exists about SSI. Those who may be entirely independent of an SSI governance arrangement have access to a level of knowledge about SSI which means they can conceive of such a project. The lesson here is that while secrecy can be of benefit if the amount of publicly available information on the topic is limited, today there may be greater obstacles to secrecy in the case of SSI than there were for nuclear weapons development at the time of the Manhattan Project.

5.5 The secrecy scenario

5.5.1 What is the secrecy scenario?

By applying Pozen's questions to the Manhattan Project, we have a clear picture of how secrecy can be conceptualised and an empirical account of how this secrecy can be realised. This will be helpful for thinking about secrecy both in theory and in relation to the practice of SSI governance. When considering the formulation of secrecy it is important to remember the level of governance which this thesis focuses on, namely global governance.

The secrecy scenario presents a shallow formulation of secrecy, whilst being sensitive to the constraint of there being a certain level of knowledge about SSI in the public domain.[158] To see that this is possible, consider whether the amount of knowledge of SSI in the public domain prevents any stage of SSI from being a shallow secret. The answer is no; the governance of SSI, its research and development, and the deployment of SSI can still be a secret. Mere knowledge about SSI does not prevent any of those stages of SSI governance from being a shallow secret, it merely prevents them from being a deep secret because the public can conceive of the possibility of SSI and the institutions to govern it.

Therefore the version of secrecy I consider now for SSI is one which focuses on the governance process. Each stage of this process would be a *shallow secret* to anyone external to this project, that is, the general public, universities, and the vast majority of people in government.[159] This is the information we need to answer Pozen's question: we know who the relevant agents are, and the variation in the level of knowledge that each group of agents has access to.

Internally, there would be extensive compartmentalisation, just like the Manhattan Project. There would be a few project leaders (the General Groves-type characters) who would be aware of most of the research and development activities which were happening in this SSI programme, but most of the employees would not be told what the true nature of the programme was. The task of researching and developing SSI would be broken down into a number of different small projects, such as understanding how long sulphur particles stay in the

[158] An account of deep secrecy for the SSI governance process would be unable to factor in the level of public knowledge that already exists about SSI. The fact that we are able to imagine such an institution shows that such an institution cannot be a deep secret.

[159] Universities may be a problematic category if certain scientists from universities are required to partake in SSI research and development..

stratosphere, what impact they have on the stratosphere, what happens when they come back into ecosystems, how they could be injected into the stratosphere, what the variations in regional impacts are, etc. These are just some examples of the different types of research questions which may be asked of SSI. The research carried out as part of these projects would take place insofar as is possible in a state of ignorance about the research being carried out for other parts of the programme.

The level of internal secrecy can be further compounded with academic journals refusing to publish the outcomes of research, as happened with the atomic energy community working on the Manhattan Project refusing to publish their work. We can conceive of a comparable practice with SSI, where the results of research projects are not published anywhere. This would dramatically limit the set of people who know the outcomes of these projects. Alternatively, we can conceive of the less restrictive practice which occurred after the Manhattan Project was completed, where researchers were allowed to publish their papers, but only in secret journals which could only be accessed by those researchers with sufficient clearance to do so. Either of these practices is a tool for increasing internal secrecy by restricting the set of actors who are able to learn about the results of research projects that are happening within the SSI programme.

Compartmentalisation combined with secrecy about research outcomes means that no single member of a research team should have the level of knowledge necessary to be able to expose the secrets of SSI. Consequently, there would be a high degree of internal secrecy.

The secrecy scenario may appear to be wired and outlandish, this could be true. Yet it is still worth considering, a reason for considering the secrecy scenario is that it explores the possibility of secrecy as a tool which could address the moral hazard. As the chapter will argue that secrecy could address the moral hazard which is valuable to know, given our reasons to be concerned with hazard.[160]

[160] Although the chapter does end up rejecting the idea of secrecy for reasons which are explained in the last section of the chapter.

5.5.2 Does the secrecy scenario address the moral hazard?

The first question to ask of secrecy scenario is whether it is sufficient for addressing the moral hazard. The answer to this question has two parts. Firstly, the answer will focus on the account of the hazard which this thesis is interested in, and then it will consider how the account of secrecy may relate to other possible formulations of the hazard. The moral hazard concern that this chapter is particularly interested in is whether the commitment of the formal role-holders of power to mitigation efforts will be adversely affected through knowledge about the research and development of SSI due to how that knowledge interacts with a set of mechanisms. I will present the answer that, as articulated in this case, secrecy does not fully address the hazard in relation to the formal role-holders of power.

Before explaining how the secrecy scenario would affect the formal role-holders of power in each state, it is important to recall the point made in section 4.3 in the previous chapter concerning the relevant agents. Namely, it is important to remember that there will be significant variation in the make-up of the formal role-holders of power across states. That is to say, some states may have very few formal role-holders of power, whilst others may have many more. For the conditions of the secrecy scenario to obtain, each state must only confide knowledge of the SSI governance regime to the minimum number of formal role-holders of power necessary.

If we take the Manhattan Project organisational tree as a map of who may need to know in the US, it appears that the President, a Secretary of State, and a Policy Committee will need to be aware of it.[161] If we were to transpose this to the case of SSI governance, then the only formal role-holders of power who would need to know from the US are the President and the relevant Secretary of State. Just two people. That is a tiny fraction of the total number of formal role-holders of power in the US, given that there at least 545 formal role-holders of power in that country.[162] This would mean that the secret-holders account for 0.37% of the formal role-holders of power in the US. Whilst other countries may for constitutional reasons have to inform more than two formal role-holders power of this governance regime, as long as the

[161] There would of course be significant variation between the organisational tree of the Manhattan Project and the global governance of SSI, given that it would be a project which was being governed at the global rather than the domestic level.

[162] As explained in chapter 4 the number 545 comes from the 535 members of Congress plus the 9 members of the Supreme Court plus the President. The number is higher if the Cabinet is also included, but the number of cabinet members can vary with each administration.

number is comparable to that of the US, it appears that the conditions of the scenario of secrecy do make SSI governance a shallow secret from the vast majority of formal role-holders of power throughout the world. Therefore, the formal role-holders of power for the most part would not experience the moral hazard at the research and development stage of SSI because they would predominately be ignorant of what was occurring.

It is worth noting that the secrecy scenario appears well placed to deal with other formulations of the moral hazard concern. Recall one of the variables of a moral hazard complaint is that of who the relevant agent is, and possible answers ranged from individuals up to international governments and corporations. Well this account of secrecy seems to minimise the likelihood of many of these sets of agents becoming aware of SSI or subject to the moral hazard. In short, most of these agents are denied the knowledge that the governance regime exists because they are not part of it. And even if certain sub-sets of them are involved, such as small teams of scientists, then their knowledge is kept to a minimum due to compartmentalisation. Hence it may be the case that they would end up working for the regime without realising it.[163] If we consider just the 'who' variable, then it would seem that the secrecy scenario is not only well placed to address the understanding of the moral hazard concern as it is used in this thesis, but also many other possible understandings of the hazard, because of the way it keeps the vast majority of possible agents ignorant of the knowledge that research and development into SSI is happening.

5.5.3 The secrecy scenario as an obstacle to procedural justice

I shall now consider a concern about the secrecy scenario preventing procedural justice. To articulate a concern of procedural justice in relation to this case, I shall consider some relevant ideas from John Rawls, Simon Caney and Marion Hourdequin. Rawls makes a useful distinction for those wishing to understand procedural justice, which is the distinction between *perfect procedural justice and imperfect procedural justice* (Rawls 1999: § 14). *According to Rawls, perfect procedural justice means that a just procedure will produce a just outcome, whilst imperfect procedural justice means that a just procedure is more likely to produce a just*

[163] As appeared to be the case for some of the employees on the Manhattan Project.

outcome, but that it is not guaranteed. Because we are engaging with a case of imperfect procedural justice, the thought is that a just procedure is more likely to produce a just outcome.

Caney elucidates the intuitive appeal of procedural justice like this:

> [T]hose who are involuntarily and profoundly affected by socio-economic forces are entitled to participate (or be represented) in the political process governing those socio-economic forces. (Caney, 2018: 3)

Caney explains that there are at least two ways of understanding the claim being made here: one could look at the 'affectedness version' or the 'subject-to-law' version of the condition (Caney, 2018: 3). The subject-to-law condition may not make sense in the context of global SSI governance, for the decision taken by such an institution may not count as 'law'.[164] Therefore, I shall focus on the affectedness version of Caney's claim.

The affectedness version is that those who are sufficiently affected deserve to participate or to be represented (Caney, 2018: 3). The affectedness principle is adopted by Hourdequin too, who holds the view that it is an important starting point which should be used to assess procedural justice for issues at the global scale, such as SSI (Hourdequin, 2019: 272–273). In the context of SSI, the affectedness principle may read as:

> *Those who are involuntarily and profoundly affected by SSI are entitled to participate (or be represented) in the political process governing SSI.*

One appeal of the affectedness principle is that it acknowledges that people are not 'mere subjects'; they have a right to influence the world in which they live even if that does not produce the best state of affairs understood in purely consequentialist terms (Caney, 2018: 4).

There is an important distinction to make when considering the possibility of procedural justice in the case of the secrecy scenario, which is the distinction between representation and participation. Depending on the account of representation which one adopts, then Caney has left open the possibility of procedural justice being achieved even in the conditions of the secrecy scenario. Substantive representation, as explained by Hanna Pitkin (1967), is understood as the situation in which an agent's best interests are represented[165] without the need

[164] Or the decisions made are not only those of creating 'law'.
[165] Pitkin offers four accounts of representation; substantive representation appears to be the most relevant for this case if we want representation to be possible.

for participation. The conditions of the secrecy scenario do not preclude the possibility of the actors in SSI governance substantively representing everyone's interests, although we may question the likelihood of this happening.

An account of procedural justice built on substantive representation as outlined in the previous paragraph would not be appropriate for the case of SSI, however. The reason for this can be seen when we consider the work of Hourdequin on procedural justice and SSI (2016, 2018, 2019). Hourdequin's argument is that given the extensive impact of SSI, we would be depriving agents of autonomy if we were to exclude these agents from participating in SSI governance (Hourdequin, 2016: 43). To elaborate, Hourdequin holds the view that the SSI governance procedure should not claim to know what the problem and the solutions are. It should not make these types of assumptions or produce framings which contain such assumptions. To do so would be to make claims about SSI on behalf of other agents, which may affect their fundamental interests. If we are going to respect these interests and create a just procedure for SSI governance, then we owe these agents participation rather than representation, on Hourdequin's account.

In the case of SSI, it is not enough that a just outcome is created for all states (although that would be impressive); there is something important about the outcome being produced by a just procedure (Hourdequin, 2019: 272). This would require a system which grants sufficient participation to agents, based on a claim such as the affectedness principle. Consequently, the procedural justice objection to the secrecy scenario is that procedural justice understood in terms of participation is incompatible with the secrecy scenario Procedural justice is an important part of justice writ large, and therefore the secrecy scenario seems to be incompatible with justice.

5.5.4 The danger of corruption

'Secrecy, being an instrument of conspiracy, ought never to be the system of a regular government.' Jeremy Bentham

Given the absence of procedural justice in the participatory sense, we know that many groups of vulnerable agents will not participate in the secrecy scenario of SSI governance. This

absence of procedural justice creates the conditions for a second objection to the secrecy scenario, namely the danger of corruption. The logic behind the corruption concern is eloquently expressed in the quote below by De Lazari-Radek and Singer, that those elites which hold power will be tempted to wield that power in a way which privileges their own interests.[166]

> Even if the lack of transparency does not lead to evils in any way comparable to those of oppressive colonial regimes, there are good grounds for objecting to dividing society into an elite and the masses. Whether it is nobles over peasants, whites over blacks, capitalists over workers, bolsheviks over the masses, or men over women, we know that those who are part of the elite will feel superior and have no difficulty in justifying, in their own terms, giving themselves privileges that in no way benefit – and often grievously harm – those they consider beneath them. (De Lazari-Radek and Singer, 2010: 53)

In the context of SSI governance, the fear is that those who wield power will do so in a way which serves their self-interest, and by doing so, they will not give due weight to the interests of those who may not be part of this process, such as the global poor, future generations and non-human nature. For this to be a genuine concern, there needs to be a danger of there being an elite group with a distinct set of interests. The fear of corruption in the context of climate change can be seen in the work of Baer *et al.* (2010), and in the context of geoengineering in work by Gardiner (2011) and Hourdequin (2016, 2018, 2019). In this sub-section I shall therefore explain some of these writers' concerns and by doing so build the argument that the secrecy scenario compounds the conditions for neglecting the interests of the vulnerable.

The fear of elites privileging their own interests is one which is well-grounded in the climate justice literature. Paul Baer *et al.* observe that the rich in poor countries may have more in common with the rich in rich countries than with the poor in their own country (Baer *et al.*, 2010: 217).[167] In context, therefore, the fear is that those who represent vulnerable countries in an SSI governance regime would focus on their own interests, which are closer to those of the people representing richer countries, as opposed to the interests of vulnerable people in their own countries.

---

[166] The fear of corruption is supported by an examination of cases of corruption in government and observing that this has not been done with the intention of other people knowing about it (Warren, 1974: 550). Infamous examples of corruption, such as the Watergate affair, begin with secrecy (*ibid.*).

[167] Of course, there are reasons which go against this as well. For example, a sense of national identity might make the rich in a particular country feel they have more in common with the poor in that country as opposed to the rich people of a different country.

The idea of power being wielded by a group of elites at the expense of other morally relevant interests is captured by Gardiner's analysis of the global and intergenerational storms. The global storm warns of the powerful nations being able to pursue policies which are in their self-interest, which would impose costs on the global poor. Yet for there to be a danger of corruption, it is necessary that the elite, in this case the powerful Western nations, have a distinct set of interests when it comes to SSI. This difference in interests can be understood if we consider that there is likely to be regional variation in the impacts of deploying SSI, which will also affect participation patterns (Robock *et al*, 2008; Tilmes *et al*, 2013; Nalam *et al*; 2018). A classic example of this is in regard to how the African monsoons may change.

Moreover, interests do not only diverge when it comes to SSI deployment. There is a plethora of SSI decisions in relation to which interests may be different. This includes where the research happens, who does the research, what information is requested of researchers, and also, very importantly, who leads the research teams and who funds the research (Smith, 2018). One may hold the view that SSI research is a neutral practice in which states do not have divergent interests, but this is far from the truth. It is unsurprising to note that different states value different things and have different priorities and concerns, and this has already been played out in relation to SSI, with some SSI researchers modelling scenarios which they claim are good for all, while researchers from other countries have observed that they cannot make such a judgement because they have not factored in how SSI may affect what is valuable to the communities in these countries, such as fishing (Hourdequin, 2019). This example sheds light on how the research stage of SSI can give rise to conflicting interests.

De Lazari-Radek and Singer articulate the fear that history is littered with elites pursuing their own self-interest, and the idea of the perfect moral storm provides reason to think that such elites also exist in the case of SSI governance. Consequently, we have reason to fear corruption in the case of SSI governance. This is compounded by an absence of publicity, which creates fertile ground for corruption to occur.

5.5.5 Moral corruption

As I already argued in the chapter on moral corruption (Chapter 3), publicity is desirable in part due to the role it can play in addressing the problem of moral corruption. To remind the reader, the argument was that a well-functioning accountability mechanism would be helpful for addressing moral corruption, and that a well-functioning accountability mechanism requires publicity so that those who hold the relevant agents in an SSI governance institution to account have access to the relevant information about that institution. As I explained in Chapter 3, publicity is concerned with the general public being aware of information and being able to access and understand that information easily.[168] Shallow secrecy requires that the information be withheld from the agent, while publicity requires that it be easily accessible by the agent. It is by definition impossible that something can simultaneously be made known to and kept secret from one and the same agent. Hence this sub-section will outline why the secrecy scenario could be undesirable because it would prevent a promising tool from addressing moral corruption.

In the chapter on moral corruption I argued that moral corruption in the case of SSI governance is undesirable because it seems to degrade the quality of the SSI governance. The reason I gave as to why this is of concern is that it increases the likelihood of unfit actors being part of the SSI governance process. Accountability is a powerful tool to address moral corruption. Given the undesirability of moral corruption, publicity and its accompanying components for addressing moral corruption are important. Consequently, there is something deeply troubling about the secrecy scenario proposal because it seems to prevent the well-functioning accountability mechanism which can address the problem of moral corruption.

Not only does the secrecy scenario prevent the use of the well-functioning accountability tactic to address moral corruption, it also creates remarkably fertile ground for moral corruption to occur. The structure of this concern is similar to the previous concern about corruption and self-interest. In the previous sub-section I made the point that corruption could occur if elites chose to pursue their perceived self-interest. In the case of moral corruption, the difference is

[168] For a more detailed account of publicity, see section 3.6 and the example from *The Hitchhiker's Guide to the Galaxy*.

that elites convince themselves that the pursuit of their own self-interest to the detriment of others is morally permissible or necessary.

5.5.6 Identifying co-benefits

There are parts of the SSI ethics literature which are implicitly supportive of well-functioning publicity. An interesting case of this is the literature on the co-benefits of SSI, an idea which is put forward by Holly Jean Buck (Buck, 2012).[169] The spirit of this idea is that deploying SSI will provide an opportunity to engage with 'human development problems such as inequality, energy poverty, food security, and land access' (Buck, 2012: 133).[170] The link between this and publicity is that if one is going to evaluate the relationship between SSI and these other human development considerations, then there needs to be sufficient knowledge about SSI and these other areas of human development, in which case the relevant aspects of SSI cannot be kept secret. For example people who research how to address food security cannot consider the impacts of SSI on food security if knowledge about the effects of SSI are a secret form them. There are variations of course in regard to the level of knowledge which can be provided about SSI. Nonetheless, if one really wants to be able to identify and understand the co-benefits of SSI, then one needs knowledge about what SSI is and its likely effects.

Some SSI researchers have taken the co-benefits of SSI into consideration. Harald Stelzer seems to be engaging in such a project when he proposes an evaluation framework to account for the normative aspects of any geoengineering technology, including SSI (Stelzer, 2017). This framework requires a consideration of how the impacts of geoengineering are linked to shared socio-economic pathways, as well as accounting for the impact, risks, feasibility and sustainability of the action (Stelzer, 2017: 150). As a result, the framework is designed in such a way that it should be able to account for all the benefits and burdens which are associated with SSI.

An analysis in this spirit is provided by the work of Heyward, who tries to assess how geoengineering technologies relate to the sustainable development goals (Heyward, 2018).[171]

[169] Buck focuses on solar radiation management (SRM) more broadly, of which SSI is one component.
[170] Joshua Horton and David Keith provide a comparable point when they argue that SSI could be used to the benefit of the global poor (Horton and Keith, 2016). This argument is considered in more detail in the following chapter.
[171] This is from a talk Heyward gave at a conference in Gratz.

Insofar as these goals provide an accurate representation of the challenges that human development faces in the near future, then it seems that this is a sensible way to understand the potential co-benefits of geoengineering for some of the greatest human development challenges that we face. Yet projects such as Heyward's also require a certain level of publicity; Heyward needed to know about the effects of geoengineering in relation to other social issues in order to be able to map the co-benefits it could offer in terms of meeting the sustainable development goals. Those who value the notion of co-benefits should therefore see the importance of publicity in being able to recognise those co-benefits. The secrecy scenario creates a significant obstacle to identifying co-benefits because it prevents the publicity condition from being fulfilled. Insofar as an understanding of co-benefits is desirable, we have reason to be opposed to conditions of the secrecy scenario in the case of SSI governance.

5.5.7 The secrecy scenario: summary

The secrecy scenario is helpful, in that it illustrates the basic but vital point for this chapter, which is that secrecy can be articulated in such a way that it appears to address the moral hazard problem. Yet such a secrecy scenario seems to be highly undesirable.[172] The often cited concern of corruption arising under conditions of secrecy does apply to this case. Moreover, it appears that the secrecy scenario presents a significant obstacle to achieving procedural justice, it is incompatible with our proposal to address moral corruption, and it limits the possibility of identifying co-benefits arising from the use of SSI.

[172] One could ask whether there are weaker formulations of secrecy which could still address the moral hazard concern yet escape the criticisms of secrecy which have been offered here. I suspect that there are weaker versions of secrecy which could address the hazard; indeed, the secrecy scenario is meant to provide an extensive account of such secrecy. However, I suspect that insofar as an account of secrecy would address our account of the moral hazard concern, then these objections would still apply to some extent, even if the strength of them was diminished.

5.6 Alternatives to Secrecy

The conclusion of this chapter may be unsatisfactory, as it has not provided a desirable way to address the moral hazard problem. It has shown how secrecy could work to address the problem, but then provided strong reasons to reject this solution. The reader may desire a positive solution to the moral hazard problem. If a simple positive solution to the moral hazard problem was available that would have been the focus of this chapter. In this part of the chapter I will comment on two other possible solutions to the hazard which are worthy of further exploration, although each raises their own problems

The first solution to be considered is inspired by Wolff's framework for thinking about decision-making about risk (Wolff, 2011, 2019). This framework was introduced in part 4.1.2 of the thesis, to show the difficulties that arise if one wishes to understand SSI as a case of moral hazard in the technical sense. Specifically, it is unclear how we are to apply the idea of an insurer-insuree relationship to the case of SSI. I argued that we do not need to resolve these difficulties, if we follow the dominant understanding of moral hazard in the geoengineering literature.

I stand by the claim that it is not necessary to understand the moral hazard in Wolff's technical sense in the case of SSI. As chapters 4 and 5 have illustrated, we can explore the problem of the moral hazard of SSI without needing it to fit the insurer-insuree model. However here I want to grant the possibility that we can understand the moral hazard of SSI according to Wolff's model, in order to explore a potential solution this opens up.

Wolff is explicit that his framework is inspired by Hermansson and Hansson's work on risk (2007). Hermansson and Hansson explain that in any risk management problem there are at least three relevant parties: the party who benefits from the risk, the party who is exposed to the risk, and the party who makes the decision about whether the risky event happens or not (Hermansson and Hansson, 2007: 132). Wolff uses this analysis to create the table below, which was discussed in chapter 2:

	Decision Maker	Benefits go to	Costs go to
Individualism	A	A	A
Paternalism	A	B	B
Moral Hazard	A	A	B
Moral Sacrifice	A	B	A
Adjudication	A	B	C

This table shows that the moral hazard problem in the technical sense can be avoided either by a change in the agents who have the decision making power, those who experience the benefits, or those who are subject to the burdens of the moral hazard. Therefore, I understand the table as pointing to four different possible paths we could take to address the hazard:

Moving Decision Making Power: give decision making power to those who are burdened by the risk-seeking behaviour, or to an agent who does not receive any benefits or burdens from the risk-taking. According to Wolff's table this means creating the structure of individualism, moral sacrifice, or adjudication.

Moving the Benefits: Move the benefits of the risk-taking to those agents who are currently burdened by it. This means creating a structure of individualism or paternalism.

Moving the Burden: Move the burdens of the risk-taking to those agents who currently benefit from it. Again this invites a structure of individualism or paternalism.

Unify the Variables: Move the variables in order that the same agent has the decision-making power, as well as receiving the benefits and burdens of the risk-taking. The only position on Wolff's table which unifies the variables is the individualism structure.

Successfully doing any of the above would mean that the moral hazard concern in the technical sense of the term is solved. Instead there would be a different structure of the relationships between decision makers, beneficiaries and burden bearers, which may come with a different set of problems. Nonetheless the moral hazard concern in the technical sense would no longer exist. However, it is not clear that the underlying substance of the moral hazard concern is addressed in any of these four options. If the core moral hazard concern is an adverse impact

of mitigation levels, then this can still occur in any of four options. These options may solve the hazard in the technical sense of the term, but it is not clear that they would address the spirit of the hazard concern. [173]

Here I will limit my comments to the issue of whether these proposed solutions fall foul of the same objections which this chapter has raised against secrecy. Recall that procedural justice in this case requires that those who are involuntary or profoundly affected by SSI participate, or are represented in, the political process governing SSI. I understand this principle as meaning those who bear benefits or burdens from SSI are entitled to representation in SSI governance. This being the case, the only procedurally just arrangement is when you *unify the variables*, by creating the individualism structure. Any other move will always leave one set of affected agents outside the decision-making procedure; for example, giving sole decision making power to the burdened would generate a complaint of procedural unfairness from those who stand to benefit from SSI.[174]

The other three objections which were raised to the secrecy proposal, those of corruption, moral corruption and inability to identify co-benefits, do not appear to apply to the four possible solutions outlined above. Each of these complaints are grounded in the effects of hiding information, which these other structures do not necessarily invite. For example, the concern about co-benefits is that secrecy deprives agents of the information about SSI necessary to identify potential co-benefits. Nothing about these other structures gives us a reason to think that such secrecy would exist.[175] This does not deny the possibility that our objections to secrecy could apply to this case, but they would apply for reasons other than that of the concealment of information.

There is one other solution I will offer, which is independent of Wolff's framework, and perhaps the most intuitive one available. This is simply that wide ranging mitigation efforts now would avoid the moral hazard problem (Barrett *et al*. 2014: 529; Baatz, 2016). If

[173] It might be easy to lose sight of the spirit of the moral hazard problem if it allows an agent to focus on addressing the problem in the technical sense of the hazard.

[174] The picture is slightly more complicated, 'justly benefit' would be a more apt description of whether an agent should be included in procedural justice. For example, we might think those who would benefit from owning slaves do not have a claim of procedural justice to take part in decisions about slavery, despite the benefit they receive from slavery. The reason for this is that the unjust nature of them owning slaves means the benefit that they receive from this practice is not just, and therefore should not have weight in any decisions about slavery. We may think that there are agents who could benefit from SSI in an unjust way, if so they would not be able to use their claim of benefit to partake in a procedurally just SSI arrangement.

[175] Although nothing ensures that secrecy would not exist in this case, but then the complaint is about secrecy not the trichotomous relationship between decision makers, beneficiaries and burden bearers.

mitigation can be ensured, then the moral hazard problem goes away; since the hazard concern is about how mitigation efforts are adversely impacted by SSI.[176]

The question then is does it seem plausible to think that mitigation efforts can be ensured? Consider a world in which a legally binding agreement exists which makes the future research and development of SSI contingent upon the meaningful mitigation efforts. This could have some force in preventing the hazard from SSI. The effectiveness of this proposal is in part dependent on whether the formal role holders of power really think that this mechanism is binding. If the formal-role holders of power think that future generations would be able to circumvent these mechanisms if mitigation efforts do not happen, due to a reason such as crisis, then they do not face that incentive to mitigate now. And can succumb to the psychological and ethical failures which induce the hazard.

To conclude this subsection, the idea of finding solutions via Wolff's table and ensuring strict mitigation efforts are both worthy of further exploration. However, they both face challenges, and run the risk of not solving the underlying substance of the hazard problem. Using Wolff's table runs the risk of only addressing the hazard in a technical sense. The strict mitigation approach would address the spirit of the hazard, although it is hard to imagine what could be done to provide confidence in such an approach.

5.7 Conclusion

This chapter set out to achieve two aims. Firstly, to investigate whether secrecy could work as a response to the moral hazard problem. Secondly, to consider whether this would be a desirable response. I have argued in this chapter that the secrecy scenario could indeed be an effective response to the moral hazard concern. However, I also argued that we have at least four reasons to find secrecy undesirable because of the impact it would have in relation to procedural justice, corruption, moral corruption and co-benefits. To make this argument the chapter applied itself to three tasks. Firstly, it drew upon Pozen's work on secrecy to answer the descriptive question of what secrecy is. Secondly, it answered the empirical question of what secrecy might look like by applying Pozen's account to the Manhattan Project. It

[176] Although the counter factual trajectory formulation of the hazard is still available in this scenario, see part 4.2 of this thesis, specifically figure 3.

considered what could be learned from the Manhattan Project in relation to secrecy in SSI governance, such as using the tool of compartmentalisation. Finally, in the normative section, the chapter discussed a set of ethical objections to the secrecy scenario

6 The intergenerational justice challenge in SSI governance

6.1 Introduction

In this chapter I seek to address the intergenerational problem of inclusion. The motive for this is that a procedurally just SSI governance arrangement would in some form include future generations.[177] Yet it is unclear what form this inclusion should take. The need for the discussion in this chapter is compounded by the complete absence of proposals to include future generations in SSI governance reports (Morrow, 2017).

Despite there being plenty of governance reports, none of them account for the inclusion of future generations. This is an indictment of these reports and an indicator of moral failure. The reports fail to take one of the most vulnerable groups (future generations) into account in their proposals for SSI governance, despite paying lip service to their relevance and vulnerability (Chhetri *et al.*, 2018, XI). Academic literature on procedural justice in SSI governance also fails to account for future generations, Callies (2019a: chapter 6) offers an account of procedural justice in for SSI governance institutions which only accounts for present generations. This is due to future generations being particularly challenging to account for in procedural justice. This chapter will therefore provide proposals for how future generations could be included in SSI governance. By doing so, the hope is that future reports will not be able to claim that there is an absence of proposals to include future generations within SSI governance arrangements.

This chapter starts by considering representation as a way to achieve inclusion. The chapter then presents the idea of *A Statement of What is Owed to the Future*, and how this could be used to help certain mechanisms which are designed to represent future generations. It will demonstrate that *A Statement of What is Owed to the Future* is helpful when it comes to applying some of the emerging tools which can be used to protect vulnerable groups, specifically: ombudspersons and constitutional protections.

Moving beyond the emerging tools for protecting future people, the main proposals presented in this chapter result from focusing on an alternative tool, that of the decision-making stages

[177] There is literature which points to the importance of including future generations in SSI decision making such as (Hourdequin, 2012, 2016, 2018, 2019; Smith 2012, 2018, Callies; 2019a). Moreover the impotence of including future generation is a plausible interpretation of what a well-designed accountability mechanism would look like, the importance of which was argued for in chapter 3.

of SSI governance and asking how future generations can be represented in this decision-making process. The chapter argues that including a second chamber in the decision-making process provides a suitable mechanism for achieving the representation of future generations in the decision-making process and thus complying with the principle of inclusion. Importantly, it will be argued that a second chamber would also benefit from *A Statement of What is Owed to the Future*. This proposal for a second chamber in the decision making stage of SSI governance is quite distinctive, given that there is general a lack of literature on second chambers for future generations, with the exception of Mackenzie (2016). Although this proposal is distinct from Mackenzie due to be it being at the international level and having a different understanding of what a second chamber should do. To clarify the scope and aims of this chapter. The chapter does not argue for a second chamber full stop. It argues that a second chamber appears to be a promising tool to address the problem of intergenerational inclusion in SSI governance. We could understand this as the chapter presenting a political theory case in favour of a second chamber. Yet this is not enough to say that we should have a second chamber in SSI governance. To make that judgement we would require knowledge of the impacts of second chambers on a much deep and interdisciplinary level. The type of knowledge which is processed by the institutional design, and international relations literature (Keohane, 1984, 1988; Fearon, 1998; Koremenos *et al*, 2001; Zürn, 2004; Raustiala, 2005; Moe, 2006; Fioretos, 2011). It is beyond the scope of this chapter to settle any debates in the institutional design literature, instead by making a case in favour for a second chamber the chapter also calls an interdisciplinary engagement with the relevant literature on institutional design, specifically to understand the implications of setting up a second chamber for SSI governance.

This chapter interacts in interesting ways with parts of the perfect moral. A perfect moral storm raises the point that our tools from political theory are bad for thinking about the problems associated with climate change (the theoretical storm), and that we face a strong temptation to pass much of the problem and harms of climate change onto future generations (the intergenerational storm). This analysis serves many functions. On one level, it is a warning about the challenge that humanity faces if it is going to overcome climate change, because it has to be able to navigate this perfect moral storm. Yet it also poses a challenge for political theorists, perhaps a call to arms, by identifying the problems that they have to be able to make sense of in their own work. This chapter can therefore be understood as a way of trying to

engage with this challenge of the perfect moral storm.[178] The inclusion of future generations is clearly relevant in terms of achieving procedural justice, yet how to do so is perplexing. Hence the problem of inclusion, which I understand to be a component of the theoretical storm.

[178] Gardiner acknowledges the importance of trying to take account of future generations in climate governance by engaging with the intergenerational and theoretical storms in his article 'A Call for a Global Constitutional Convention Focused on Future Generations' (Gardiner, 2014: 305).

6.2 How should future generations be included: participation or representation?

A principle of inclusion could be realised in at least two ways. One is for the group in question to meaningfully participate in SSI governance; the other is for it to be well represented in SSI governance. In this section I will consider the possibilities of participation and representation.

We have reasons to think that a principle of participation is preferable, for example there is work in the literature on ethics of SSI which makes the point that vulnerable agents ought to be recognised in SSI governance (Hourdequin; 2016; 2018; 2019). Recognition in turn requires those agents to be able to engage with the governance process and exercise their voice. It is clear what the problem is though if we are thinking of future generations, they are yet to have a voice to exercise. Given the voicelessness of future generations it is clear that participation would not be a plausible way to include future people.

Representation does not necessarily require the voice which participation does.[179] Although there is still a challenge if one wishes to think of representation as a means to achieve inclusion. Which is what is meant to be represented? When we think of representative systems we might think that one of at least two functions are being achieved by this representation. We may think that we are representing preferences or representing interests of the relevant agents.

Yet when thinking of the interests or preferences of future generation we have an epistemic problem (Karenin, 2016: 86). This refers to our inability to know the specific preferences and interests of future people. The epistemic problem is exacerbated by the problem of plurality (Karenin, 2016: 87), which observes that across time and space there will be a great variation in preferences. It strikes me that these features make it remarkably challenging to anyone who would wish to try to take the preferences of future generations into account.

Yet this problem is not insurmountable, nor even particularly problematic for certain interest-based accounts of engaging with the problem of voicelessness. The reason for this is that certain interests are not contingent upon the particular person in question because they exist simply in

[179] This may be part of the reason why representing future generations has started it taken seriously in the literature about institutions which are meant to help climate justice be achieved (Gardiner, 2014; Lawrence, 2014; Stevenson & Dryzek, 2014; Heyward & Roser, 2016). Additionally there is an understanding that the representation of future generations is a desirable way to achieve both climate and intergenerational justice (Van Parijs, 1998; McKinnon, 2011; Roser & Seidel, 2017; Lawrence & Kohler, 2018).

virtue of being human (Karenin, 2016: 86).[180] Given that we have the epistemic power to know the specific interests of future people insofar as they are human, the question then becomes one of what can be done to represent those interests in the shared politics of the current era. In the absence of future generations being able to represent their own interests, I propose as the next best alternative that those interests be represented in a meaningful way by the present generation. This move seems to be supported in the literature on future people which, whilst it does consider representation, does not (insofar as I know) take the idea of the participation of future people seriously.

The problem of voicelessness is the relevant obstacle to the participation of future generations under discussion in this chapter. Therefore I shall proceed by proposing ways to address the problem of voicelessness for future generations in SSI governance. This is a lower bar than that of participation, which seems to demand the presence of the agent in order for them to exercise their voice and reason.

[180] A similar logic has been used by Shue (1980), and I take Simon Caney to be doing something similar in his chapter 'Climate Change, Human Rights, and Moral Thresholds' (2010), when he chooses a minimal account of human rights and demonstrates that such an account is sufficient to show that those human rights are threatened by climate change.

6.3 A good minimum: *A Statement of What is Owed to the Future*

In this section I shall introduce an integral part of the strategy which this chapter offers for addressing the problem of intergenerational inclusion. The idea is that there should be a statement of *what is owed to future generations*. Such a statement should not be seen as a definitive account of all things that are owed to future people. It is a statement of a *minimum* which is owed to them. Such a statement should therefore outline a minimal account of the interests of future generations.[181] The reason for being interested in this statement is that it could be used to improve the functioning of mechanisms which are designed to account for or protect the interests of future generations. To consider this proposal, the section will do two things. Firstly, it will explain what the statement is, and secondly, it will consider how the statement relates to another important proposal regarding future generations, that being Gardiner's Global Constitutional Convention for future generations (Gardiner, 2014). The motive for considering this is that *A Statement of What is Owed to the Future* could be a useful part of the strategy to enable their inclusion in an SSI governance institution.

6.3.1 The declaration

The statement, or declaration, of what is owed to the future will be a document which outlines what future generations are owed by the present generation.[182] The declaration can be understood as functioning in the way declarations have done in the past when they are designed to protect fundamental interests. Notable examples of this include the UN Declarations of Human Rights, the UN Declaration on the Rights of the Child, and the UN Declaration on the Rights of Indigenous Peoples. The *Statement of What is Owed to the Future* needs to be

[181] I use the term interests, which loads such a statement towards a conversation about the rights of future people, due to interests providing the foundation for prominent rights theory. If rights are not the appropriate metric of what is owed to future people, then it may well be the case that the language of interests is inappropriate for such a statement. However, it appears that there is no neutral term for a metric of what is owed to future people. Added to which, the language of interests has the advantage of avoiding non-identity-related problems, insofar as a set of interests can be held by any possible set of future people, regardless of their specific identity. The reason for interests loading the conversation in the direction of rights is that interests can be a strong source of rights.

[182] Owed is being used broadly here; one may prefer to understand this as entitlements or rights of future generations.

supported by the vast majority of states in the same way as these other examples which have been adopted by the UN General Assembly.

The content of the declaration would be determined by those who draft and sign it. This is in order to be sensitive to the way in which declarations generally come into existence.[183] Yet this creates the risk that the statement would not be even minimally just. For example, the statement could proclaim that nothing at all is owed to future people. I shall assume that the declaration which is agreed to will be minimally just. This assumption is not a probabilistic claim that such a statement is *likely* to be minimally just;[184] it is based on the premise that the problem of intergenerational inclusion is a serious one, and consequently it is worth considering the more promising strategies to address it.[185]

I shall now consider what a minimally just declaration may look like. This is helpful insofar as it provides an example of the core content of the declaration, which we can take to be illustrative of any declaration of what is owed to future generations which might be agreed to by states. To consider what the core content of the declaration may look like, I shall draw on an article by Simon Caney in which he outlines how a minimally just account of human rights would apply to the case of climate change (Caney, 2010).[186] In the article he argues that if we accept a minimal account of human rights – i.e. life, health and subsistence – we are able to make substantial demands in the context of climate change, given that climate change threatens these minimal rights (Caney, 2010: 166–169).

I propose that the declaration of what is owed to the future could adopt a Caney-like minimal account for future generations. In other words, that future generations have the right to life, health and subsistence. For the purposes of this chapter I shall understand part of the core content of the declaration as a version of Caney's minimal account of human rights oriented towards future generations. Keeping in mind nonetheless that the content will not be determined by philosophers, but by those who actually write and sign the treaty.

---

[183] There may also be normative reasons as to why it is important for the declaration to be drafted by these law-makers, perhaps on the grounds of legitimacy.

[184] On the contrary, if one accepts the usefulness of the statement, then a serious question is under what conditions a just version of this statement is more likely to occur.

[185] Whilst it may be an interesting question whether an unjust declaration of what is owed to the future could be of use for addressing the problem of intergenerational inclusion, it is not the question which this chapter will explore.

[186] This bears a relation to Henry Shue's account of basic rights, in which he outlines a minimal set of rights which need to be realised if any other rights are going to be protected.

The benefits of such a minimally just declaration are at least threefold. Firstly, such a minimum should be reasonably acceptable to many reasonable persons, for it is not contingent upon a specific sophisticated account of intergenerational justice, but merely what is *at the least* owed to future people on many accounts of intergenerational justice. This means the acceptability of the statement should not hinge on a position one adopts about the currency of intergenerational justice. This is important due to the variety of currencies available for intergenerational justice and a lack of agreement about which the appropriate one is. The classic currencies of distributive justice, such as welfare (Page, 1983), resources (Barry, 1999, 2005), and capabilities (Page, 2007; Gutwald *et al*, 2014) have all been applied to the intergenerational context.[187] Determining the appropriate currency is an unenviable task, a strength of this statement is that statement should not be contingent upon an answer to this question.

Secondly, such a minimum, if met, ought to provide a sufficient state of affairs for future generations. This does not necessarily mean that the state of affairs would be just or optimal for future people, but that they would live in a world in which at least their basic interests were fulfilled.[188] This is similar to how if everyone's basic interests were protected today, then there would be no extreme poverty. It does not mean justice has been achieved (assuming that the content of justice is greater than the meeting of basic interests), but it is also good in the same sense that no one is subject to the harm of extreme poverty.

The third benefit of such a minimal account is that the statement should be applicable to those institutions which are meant to engage with the policy that affects these interests. This will become much clearer once I have applied this statement to the specific proposals for decision-making procedures which can give weight to the interests of future generations. Although there will of course be challenges in applying the principles from this statement to public policy. The spirit of these three points is that a minimal account ought to be acceptable to many people, and it will still be helpful for protecting future generations.

[187] For a good overview of the currencies of intergenerational justice see Page (2007) as well as Vrousalis (2016), the issue is also covered briefly in Caney (2018). For a critique of the egalitarian currencies of intergenerational justice see Beckerman (1999).

[188] Of course, there are questions about how sufficiency is understood in relation to justice, and for some, the measure of justice for future people may well be that of sufficiency (Benbaji, 2005; Brown, 2005; Casal, 2007: 297–298). Arguably, Rawls also holds a sufficiency view with his account of just savings.

6.3.2 *A Statement of What is Owed to the Future* as complementary to a Global Constitutional Convention Focused on Future Generations

Such a statement can be understood as constituting part of what is called for by Gardiner in his article 'A Call for a Global Constitutional Convention Focused on Future Generations'. The aim of such a convention is to provide a 'constitutional system that appropriately embodies intergenerational concern' at the international level (Gardiner, 2014: 305). The scope of the statement that I am proposing is the same as Gardiner's, as it should outline the interests that can be applied to any international governance regime that raises intergenerational concerns (although this chapter will only focus on the application of the statement to SSI governance).[189]

The *Statement of What is Owed to the Future* could be understood as a component of the Convention, which will provide 'a set of norms (rules, principles or values) creating structuring and possibly defining the limits of government power or authority' (Gardiner, 2014: 306). It is unclear how demanding it would be to create this set of norms. Gardiner does mention that this should be understood in a minimalist sense, in which case this may effectively be the same thing as *A Statement of What is Owed to the Future.* For the statement in effect places a restriction on what governments can do by stating what they should not violate. Yet the creation of rules, principles or values may be much more demanding than what *A Statement of What is Owed to the Future* can achieve. Given the ambiguity of what is being called for by Gardiner in this instance, I presume there to be a benefit in calling for an explicit *Statement of What is Owed to the Future,* which is the type of document that a Global Constitutional Convention ought to produce.

Gardiner explains two attractive features of a Global Constitutional Convention. Firstly, the Convention is based in political reality. By this Gardiner means that the Convention understands the relevant issue at hand, i.e. climate change, and the reasons why we have failed to respond to it (as explained in Gardiner's perfect moral storm), and that an alternative system is required to address intergenerational problems (Gardiner, 2014: 306). The second attractive feature is that the Constitutional Convention does not respond to political reality by entering a state of despair (Gardiner, 2014: 306). It uses the political reality in the correct way, and by

[189] It is interesting to note that although the statement is motivated to engage with the idea of including future people in regard to SSI governance, because it provides a minimal account of what is owed to future people, it might be expected that it would also be applicable to cases other than SSI (such as food, heath and sustainability), due to future generations having minimal interests which are applicable in other policy areas as well.

understanding the problem at hand, the Convention can then proceed in such a way that it is sensitive to the nature of the problem. I understand this account of political reality as emphasising the need for responses to a problem like intergenerational inclusion as being fit for purpose. This provides an additional reason for thinking that the statement at least needs to be minimally just, if it fails to be minimally just it would appear not to take the nature of the problem seriously.

It is worth distinguishing between political reality and more common but similar sounding terms, such as political feasibility. Political feasibility is a concept which is used to rule out proposals which cannot be implemented, and to rank proposals according to the likelihood of them being implemented (Gilabert and Lawford-Smith, 2012: 815). Taking political reality seriously could sound as if one taken feasibility constraint very seriously. Importantly this not the case, taking political reality seriously does not require one to consider the likelihood of whether the response can be implemented, but whether the response is sufficient given the nature of the problem.

A Statement of What is Owed to the Future should be able to retain these two attractive features of Gardiner's Global Constitutional Convention. Firstly it does not respond to the problem of intergenerational inclusion by entering a state of despair instead there is the creation of *A Statement of What is Owed to the Future* with the goal of helping mechanisms function to represent future generations. This Statement shall also be grounded in political reality. I understand this as the statement being sensitive to the obstacles which are presented in order for it to be fulfilled. Consequently the statement should draw attention to the demandingness of its content and how the present global system fails to achieve what the statement requires. By doing so the statement sets up the challenge which is meant to be addressed.

It should be noted that this section does not provide a strong defence of the *Statement of What is Owed to the Future*. Instead, it has pointed to the benefits of setting a minimum standard and how this relates to Gardiner's Global Constitutional Convention. The case in favour of this statement will become clear throughout the chapter. The case is grounded in how different proposals for the inclusion of future generations benefit from *A Statement of What is Owed to the Future*. The chapter will draw attention to these benefits after analysing different proposals to promote the inclusion of future generations within SSI governance.

6.4 Applying *A Statement of What is Owed to the Future* to the tools which are meant to protect future generations

A Statement of What is Owed to the Future could work in conjunction with some powerful tools for taking future generations into account. These tools comprise the use of ombudspersons and constitutional protections. The reason for looking at these particular tools is that they have been vividly defended in the contemporary literature on designing institutions to protect future generations (Gonzalez-Ricoy and Gosseries, 2016).[190] Moreover, the idea of having ombudspersons for future generations has already been practised in Hungary (Gonzalez-Ricoy and Gosseries, 2016: 4). Similar practices have been adopted elsewhere, such as in Israel, where there is a commissioner for future generations, and in Finland, which has a committee for the future. I will now proceed to illustrate how *A Statement of What is Owed to the Future* is able to facilitate the use of these tools.

6.4.1 Constitutional protections

Constitutional protections for future generations are advocated by Iñigo González-Ricoy (2016). The idea is that certain interests of future generations can be given legal protection within the framework of a constitution. González-Ricoy argues that constitutional protections have at least three advantages. Firstly, they reduce short-termism by creating a cost for policy-makers if they try to remove these protections (González-Ricoy, 2016: 173). Secondly, they can reduce long-term uncertainty: if a constitution is designed to protect long-term projects such as protecting the future then agents may have more faith in those projects (González-Ricoy, 2016: 176).[191] And thirdly, constitutions can be informative in regard to society's values

[190] Although there are also other noteworthy proposals to protect future generations, broadly speaking these proposals either focus on ensuring that office holders are more future friendly, or that office holders are have greater incentives to be future friendly (Gonzalez-Ricoy and Gosseries, 2016: 8-9). Examples of ensuring that office holder are more future friendly include ideas such as having some seats in a parliament put aside of membership of environmentally friendly groups (Dobson, 1996), or youth quotas for a parliament (Bidadanure, 2016). This is in contrast to providing incentives, or mechanisms for office holders to be more future friendly, for example Caney among a set of proposals argues that governments should provide a manifesto of the future (Caney, 2016: 136).

[191] As opposed to long-term projects being pursued when there are no long-term constitutional protections.

(González-Ricoy, 2016: 178).[192] This can happen by constitutions challenging its citizens to reflect upon the importance of the values which are entrenched within it (González-Ricoy, 2016: 178).[193]

However, González-Ricoy also highlights that there are reasons to be concerned about constitutional protections, not least the uncertainty about whether we may harm future people with our choice of constitutional protections (González-Ricoy, 2016: 171). This is an old concern about constitutions, on which Thomas Jefferson was particularly vocal (Jefferson 1999 (1789), although there is plenty of recent literature which tries to dampen the force of this concern (Gosseries, 2008, 2014)

Yet there is a more immediate concern with such constitutional protections when it comes to global SSI governance. The problem is that there is no global constitution, and therefore there is no appropriate framework that the interests of future generations can be protected within. However, this need not be the case, and González-Ricoy explains that three conditions need to be met for constitutional entrenchment to happen:

> First, they are included in a legal document (i.e. the constitution) with normative superiority over ordinary statutes… Second, they can only be amended by means that are more stringent than those of ordinary law-making procedures… Third, they are enforceable by some independent body (typically a constitutional court). (Iñigo González-Ricoy, 2016: 172)

If González-Ricoy has identified the relevant criteria, then it appears that constitutional entrenchment is possible at the global level, even in the absence of a global constitution. If we consider the formation of a global SSI governance regime, then such a regime can be founded on a document which has superiority over the documents that such a regime can produce. The process for amending the founding document can be made more challenging than that for ordinary decisions taken by the governance regime. Finally, this document can be designed in such a way that it is enforceable in international courts. The possibility of the founding document of SSI governance having this constitutionally entrenched character also raises a set

[192] America seems to offer a clear example of a case where many agents identify with particular values within the constitution, such as the value of free speech or the right to bear arms.
[193] The good of this would in part be contingent upon the desirability of those values. In this case thinking about the value of future generations appears to be a worthwhile value.

of legal questions about the document concerning its legal status and how it would relate to international courts.[194]

We can understand the relationship between constitutional protections and the statement of what is owed to the future in at least two ways. The first is that the statement could be informative of the constitutional protections which future generations are granted. The second is it could be considered a constitutional protection in itself .The answer to this is contingent upon the status of the statement. I will now outline the logic of these two positions..

If we accept the possibility of putting constitutional protections in place for future generations, then there is still the question of what should be the content of these protections. *A Statement of What is Owed to the Future* may be helpful in answering this question. As such a statement provides an account of the minimal interests of future people, it contains much of the information that is required for constitutional protections to be formulated. The question then becomes one of what are the relevant interests from this statement which ought to be protected via a constitution. If the interests which the statement takes into account are genuinely the minimal ones, then we can expect the constitution to include all of these interests in order to be considered minimally just. Similarly, if there are any weighty additional interests which are not listed in *A Statement of What is Owed to the Future,* then we can expect a debate on whether those interests should be protected. Therefore the scope of interests accounted for in constitutional protections may be greater than *A Statement of What is Owed to the Future*

The statement can also be understood as a form of constitutional protection. To do so would require that the statement meets the three conditions set out in the previous page: it would need to have greater force than an ordinary statute, it would have to be harder to amend, and it would have to be enforceable. Constitutional protections require enforcement mechanisms. If we were to understand the statement as form of constitutional protection we have an additional reason to be concerned with what mechanism could guard the content of the statement.

[194] These matters would have to be explored in much more detail if this option were adopted.

6.4.2 Ombudspersons

An ombudsperson's concern is to ensure compliance with whatever rules, treaties and laws exist (Beckman and Uggla, 2016: 126).[195] In order for an ombudsperson to carry out their function, they require some account of the relevant rules to which they are holding an institution to account. *A Statement of What is Owed to the Future* can provide such an expression of the rules.[196] If *A Statement of What is Owed to the Future* sets out the rights of future generations, then an ombudsperson should have the power to act when the SSI governing body appears to violate these rights. *A Statement of What is Owed to the Future* gives an ombudsperson for future generations the information that they require to function.

Yet ombudspersons lack formal power and cannot enforce compliance with any of their recommendations (Beckman and Uggla, 2016: 124). It is tempting to understand this as a disadvantage of the ombudsperson system. After all, the interests of future generations are very significant, and it may appear that we are underestimating the seriousness of the problem if we task the protection of these interests to an actor with no formal power.

However, such a view fails to account for certain advantages of an ombudsperson, because the lack of formal power allows them to avoid concerns about democratic legitimacy (Beckman and Uggla, 2016: 124). The legitimacy argument states that if the ombudsperson could enforce compliance then there would be a problem as to whether the ombudsperson would be considered legitimate, given that they are not elected and exercise power over elected representatives (Beckman and Uggla, 2016: 124). The prospect of an ombudsperson may seem unsatisfactory: if an ombudsperson is not a constraint on power, then what is the point? Here the literature teaches that ombudspersons can be a constraint on power even if they lack formal power, by generating popular support for their work from the electorate, to which politicians in turn respond (MacKenzie, 2016: 35).[197]

---

[195] Beckerman and Uggla note that the idea of an ombudsperson for the future has been vigorously defended, citing Weiss 1992, Slaughter 1994, Aguis 2006, Shoham and Lamay 2006, Javor 2006, Gosseries 2008 and Gopel 2012 (Beckman and Uggla, 2016: 117).

[196] Rules being understood in a loose sense of the term, due to the fact that ombudspersons do not require extensively detailed legislation; they merely require legally valid documents on the basis of which to ask questions (Beckman and Uggla, 2016: 126).

[197] This popular support creates pressure for those with power to be sensitive to the judgement of the ombudsperson (MacKenzie, 2016: 35).

It is worth noting that 'ombudsperson' may not be the appropriate name for an international actor (such as one who might act within the type of SSI governance institution which we are considering), but an equivalent is surely capable of being created.[198] A comparable set of actors within the international system is that of the UN Special Rapporteurs. These are independent actors who have the relevant expertise to report on a special issue within a given mandate (Freedman, 2014: 112).

Yet there are important differences between ombudspersons and UN Special Rapporteurs. For example, rapporteurs are subject to strict term limits, which in total cannot exceed six years. This does not seem to be the case for ombudspersons, and in fact, one of their appeals is that they are able to think over the longer term. A term limit would go against the functioning of an ombudsperson if it encouraged them to adopt more short-term interests. Additionally, an ombudsperson performs a function which is closer to that of mediation, whilst a rapporteur's role is purely advisory. If a special rapporteur is an inappropriate equivalent for an ombudsperson, then this just creates an opportunity for SSI global governance to contribute to the development of global governance more generally, by pioneering the international ombudsperson.

The question which will permeate this chapter is what mechanism we should employ to enact the statement of what is owed to future generations. Ombudspersons are a plausible way of trying to uphold the statement of what is owed to future generations. Yet they are not without their shortcomings. I shall draw attention to an important shortcoming of an ombudsperson as a way to protect future generations, and then consider a second chamber as a mechanism which can protect the content of a statement of what is owed to future generations, whilst not suffering some of the important shortcomings of an ombudsperson.

Ombudspersons do not appear to have a strong influence at the time at which a decision is made. They exercise their power after a decision has been made. Considerations of procedural justice remind us that, insofar as it is possible, we should represent the interests of future people when decisions are being made which affect them. [199] There is a simple instrumental reason to think that this is important, which is that if the interests of future people are properly represented then decisions should not be made which are detrimental to these interests. The absence of the ombudsperson from the decision-making procedure opens up the door for

[198] The role of the ombudsperson appears to exist at the national level only.
[199] For more details on procedural justice see part 5.5.3 of this thesis.

decisions which are harmful to future generations. Whilst the ombudsperson may eventually help this decision be overturned, it would be preferable if these decisions were not made in the first place.

One may think the difference between a second chamber an ombudsperson is not too important, if both eventually reject unjust decisions towards future generations, why would it matter if this rejections happens in a second chamber or by an ombudsperson. One answer is that an ombudsperson may not be able to reverse all the impacts of a future effecting harmful decision.

It is desirable that the mechanisms for protecting *A Statement of What is Owed to the Future* is consistent with the demands of procedural justice. Therefore, I want to find a mechanism which represents the interests of future generations at the time at which a decision is being made.

This should not be understood as a claim that second chambers are always better than ombudspersons, or that ombudspersons are procedurally unjust. Instead the claim is that second chambers are better at meeting the demands of procedural justice for future generations in the case of SSI governance. Yet there will be future effecting decision which do not require a second chamber to be considered procedurally just, for example decisions made by local councils about infrastructure projects. In this case it may be sufficient for there to be an ombudsperson to whom complaints can be made in the grounds of how future generations are effected by the decision.[200]

Yet there will be cases where considerations of procedural justice are not weighty, and ombudsperson would be more appropriate. For example in the UK there is an ombudsperson for complaints about communication companies. It would be odd to think that a second chamber would be required due to considerations of procedural justice to make sense of communication decisions instead.

To summarise this section: *A Statement of What is Owed to the Future* should be easily accessible and understandable. Specifically, such a statement should be beneficial for the functioning of some of the most prominent tools available for protecting the interests of future people: ombudspersons and constitutional protections. In addition, *A Statement of What is*

[200] The reason why the demands of procedural justice vary can in part be understood as a product of the weight and relevance of the interest which the effected parties have. In SSI governance the interests of future generations are very strong given that there minimal rights can be threatened, the weight of interest demand that procedural takes these interests sufficiently serious. In contrast infrastructure projects do not appear to threaten the rights of future generations, the stakes for future generations are not as high in these decisions, , hence an ombudsperson may be more appropriate for representing the interests given there relatively lack of importance.

Owed to the Future can be understood as being complementary to Gardiner's Global Constitutional Convention. Moreover we have reasons to think that a second chamber is more desirable than ombudspersons in this case due to their ability to come close to meeting the demands of procedural justice.

6.5 A second chamber? Giving weight to the interests of future generations

The previous section considered some prominent tools for the inclusion of future generations in SSI governance. I will now consider a different tool for how the interests of future people can be accounted for in SSI governance. Therefore, the focus now moves away from the emerging tools for protecting the interests of future generations to a consideration of tools for including future generations when the decision is actually being made. Specifically, I am going to focus on the decision-making procedures which occur within some form of a decision-making chamber. (Some people may be more comfortable with alternative language to describe this chamber, such as a decision-making body or organ, which is the preferred language used in certain UN institutions). I will consider how such a chamber ought to be designed to take the interests of future people into account.

6.5.1 A second decision-making chamber

The proposal which I will consider for giving weight to the interests of future people is that of a second decision-making chamber.[201] It is not uncommon for a second chamber to be charged with considering a different set of interests from those of the primary decision-making chamber, in the hope of improving the quality of legislation which is produced from the decision-making process as a whole. The thought here is that a second chamber could be charged with specifically protecting the interests of future generations.[202] In this sub-section, the main consideration is how a second chamber should be formulated. To do this we will engage with the work of MacKenzie (2016), who argues in favour of three principles to determine the composition of second chambers at the national level which are meant to protect future generations.

[201] Including future generations at the point in time in which the decision is made is one of distinctive advantages of a second chamber from the other tools which have been considered. The consideration of a second chamber in this context is also interesting for there is a distinct lack of literature on how second chambers can represent future generations, although there is broader on how institutions can be incentive to account for future generations (Caney, 2016; Ekeli 2016; Iñigo González-Ricoy; 2016).

[202] This is in contrast to the primary chamber where the initial decision within the SSI governance process occurs. More will be said later about the relation between these two chambers.

Prior to exploring the benefits of a second chamber I will first provide the reader with the necessary tools to make sense of chambers at the global level of SSI governance. As was explained in section 1.3.2 of the introduction, the level of governance that this thesis focuses on is the global level. Here I wish to draw attention to a particular aspect of my understanding of global governance, which is that there has to be a space where decisions about SSI are formally made.

This space would be comparable to the formal decision-making space which exists in many global institutions such as the UN General Assembly, the UN Security Council and the UN Economic and Social Council, as well as international organisations outside the UN such as the International Monetary Fund, the World Bank and the World Trade Organisation: there is a process within each of these organs where decisions are formally made. Considering the UN General Assembly, for example, representatives of all the member states come together for debate, each member state has one vote, and depending on the decision at hand either a simple majority or a super majority will be required for the decision to be approved. A detailed articulation of how those organs function is not necessary for our analysis, since the analysis is not contingent upon a detailed picture of the function of a primary chamber. It merely requires an understanding that there can be formal spaces in the international arena where decisions are made on particular topics. Given that our focus is on SSI governance, I propose we understand a primary governance chamber as a space in the international arena where the vast majority of states come together to make decisions about SSI.[203]

A second chamber is a separate space to that where the primary decisions are initially made in regard to SSI governance (such as an assembly of states, although there is variation concerning what this process would actually look like). Decisions made by the primary chamber would then be subject to scrutiny by a second chamber. I envisage this second chamber as having two distinct sets of powers in regard to the decision-making. Firstly, it should be able to propose amendments to decisions made by the primary chamber.[204] The nature of these amendments would be restricted to those which can be justified as protecting the relevant interests outlined in *A Statement of What is Owed to the Future*. The second potential set of powers would relate

[203] If the idea of this type of global governance of SSI were to be advocated for there would be complicated questions around how such an institution would relate to other global governance bodies such as the United Nations Environment Programme (UNEP), which may seek to claim some jurisdiction over the governance of SSI. Such questions are put to one side for the analysis in this chapter, which is only interested in the potential benefits of a second chamber, not the legal relationship between potential chambers and current international governance regimes.

[204] This is in the spirit of the power which the UK's House of Lords has in the legislative process.

to preventing a decision being passed by being able to challenge it in an international court of law.[205] This power has been proposed by Kristian Ekeli, although rather than suggesting the establishment of a second chamber per se he proposes that this power be given to minorities within a chamber (Ekeli, 2016).

Here I shall comment on the balance of power between the second and primary chamber, in order to make the idea of a second chamber clearer. A simple understanding of the balance of power between the primary decision-making chamber and the second chamber is that the second is a constraint on how the primary chamber can behave in regards to how its actions affect future generations. It wields a veto power over decisions, which prevents the first chamber violating the interests expressed in *A Statement of What is Owed to the Future*. This veto power may make the second chamber sound very powerful, but it can only wield that power in very specific circumstances, namely when the primary chamber violates the interests set out in *A Statement of What is Owed to the Future*. There are further details to be determined about the balance of power between these chambers, such as should the second chamber have any greater power than that of a veto in specific circumstances? The appropriate balance of power between chambers and other institutions is a concern to be determined by experts of institutional design in consultation with others to work out the desirable balance

MacKenzie (2016) argues that a randomly selected second chamber could be well placed to consider many intergenerational concerns. Helpfully, he provides three criteria for addressing the 'short-term tendencies of elected legislatures' (MacKenzie, 2016: 285). I shall now explore MacKenzie's criteria and consider how they could be applied to the case of a second chamber for SSI governance.

The first criterion is one of independence (MacKenzie, 2016: 285). For MacKenzie, independence appears to mean that the representatives need to be independent of the political pressures which induce short-term thinking. That is to say, they are not subject to the electoral pressure which drives short-term thinking and action. If independence does relieve the pressure of short-term thinking, then that appears to be a sensible condition to have for a chamber which is charged with defending the interests of future generations.

In the context of SSI global governance I agree with Mackenzie's first criterion of independence. MacKenzie expresses this by arguing that the participants in the institution

[205] The question of which international court would be appropriate is one for scholars in other fields to answer, probably legal scholars.

should not be elected by the people. This coheres with the common practice of global governance institutions where no representatives are elected by the people. This should mean that these representatives would not suffer from the short-term incentives of an electoral cycle.[206] Nonetheless, other provisions may also be required to help representatives to be truly independent; for example, there is a danger that representatives might be subject to other short-term pressures, such as their job being contingent on short-term gains for the country that they represent.

MacKenzie's second criterion is that of representativeness. By this, MacKenzie means that the chamber should represent the full range of interests which exist within the general population (MacKenzie, 2016: 285–286).[207] The particular interests that MacKenzie seems to have in mind are those of where different time horizons can be adopted.[208] There is a question about how representativeness should be understood in the context of a second chamber for SSI governance. The function of representativeness for MacKenzie is that the full range of interests should be accounted for and properly represented.

The proposed chamber does not meet Mackenzie's stand of representativeness because a full range of interests are not represented within it. Instead a core subset of full range of interest is represented within this chamber. Perhaps an ideal second chamber would represent a full range of interests of future generations, yet the epistemic barriers to knowing this full range, a minimal core set of interests has an appeal.

The third criterion which MacKenzie highlights is that of deliberativeness. As he argues, this is a way of improving long-term thinking thanks to the provision of information about the impacts of long-term action and how actions are interrelated and therefore can have an impact upon the long term. This feature appears to be helpful, given that the influence of SSI decisions on the interests of future people is complicated and not easily understood. A well-structured

[206] One could fear that the institution would not be legitimate if participants are not elected (Bodansky, 1999). This type of concern is addressed in work on global governance institutions which argues that they can be legitimate even if the participants are not elected, as long as certain conditions are meet such as them being accountable (Grant and Keohane, 2005; Buchanan and Keohane, 2006).

[207] Mackenzie's account of representativeness can be a bit confusing insofar as much of it is focused on how demographic representativeness is good for achieving representativeness of time horizons, for the proposes of my account I am focusing on the his end goal of representativeness of time-horizons as oppose to his mean of achieving this via demographic representation.

[208]In context time horizon means the period of time which an agent is interested in. For example a short term time horizon on climate change would be to consider its impact in the coming year. Whilst a long term time horizon would consider the impacts of climate change in 50 years time.

deliberative process seems desirable in order to have a clearer understanding of how a decision would affect the interests of future people.

I accept MacKenzie's point that deliberativeness is important for understanding the impacts of any decision on future generations. He points to the usefulness of providing the relevant information for a deliberative process to function well. I have little to add here. A well-functioning deliberative process clearly seems to be able to help second chambers account for the interests of future people. The question is one of how such a process could be designed, which is not one I address in this thesis.

The second chamber should cohere with MacKenzie's standards of independence, representativeness and deliberativeness. As explained, the powers of this chamber would range from being able to amend decisions made in the primary chamber to being able to challenge the decisions of the primary chamber in an appropriate court of law. Having presented an image of how a second chamber could be understood in this context, the question now is why we should find it desirable to have a second chamber. The following sub-section will therefore offer a defence of this proposal for a second chamber.

6.5.2 A defence of second chambers, considering the arguments from history

The previous sub-section articulated some of the core features of a second chamber. I shall now defend the focus on second chambers. Second chambers have an extensive history and they are arguably as old as democracy, with second chambers being used in ancient Greece and the Roman Republic (Shell, 2001: 6). Many arguments have been offered in favour of second chambers, and here I shall outline the view of three pre-eminent thinkers on the topic, namely: Aristotle, Madison and Mill. The arguments they make are not designed for a second chamber at the global level of governance, however. Therefore, once their arguments have been considered, the case will not yet have been made that these arguments are necessarily relevant to the evaluation of a second chamber at the global level.

Aristotle defended the ancient Greek system of a 'mixed government' by arguing that representing different interests would help society remain stable (Aristotle: IV–V; Biondi,

2007: 189; Russell, 2001: 19).[209] In contrast to this view, John Stuart Mill argues in favour of a second chamber which should be composed of states-people who lack any class interest or prejudice (Mill, ([1861] 1968: chapter 13). This is very different to Aristotle's idea that the second chamber should represent particular classes and it is fun to see how such contrasting ideas can both lead to support for a second chamber.

James Madison, the fourth president of the United States, was also in favour of two chambers, arguing at the Constitutional Convention in Philadelphia that the role of the second chamber was 'first, to protect the people against their rulers, second, to protect the people against the transient impressions into which they themselves might be led' (quoted in Patterson and Mughan, 1999: 14). The theme of Madison's thinking is that second chambers are protective chambers for a particular set of people, and this will be revisited in the next paragraphs. Whilst Madison had chambers at the national level in mind,[210] his argument applies to a second chamber at any level of governance.

There is nothing special about national-level governance to mean that the national level is the only level at which second chambers are appropriate. The appropriateness of having a second chamber at any level is contingent on the type of power it has in relation to the primary chamber. Both Mill and Madison appear to have similar motivations for finding a second chamber desirable; namely, they both think that it improves the quality of decision-making. Mill holds this view in part due to this thought that a second chamber will be composed members who are devoid of class interests or prejudices, for Mill this helps the members of a second chamber be good states-people (Mill, 1861: chapter 13). By this I take him to be making the same point as Madison when Madison argues that a second chamber should 'protect the people against the transient impressions into which they themselves might be led'. A way of doing this is to have good states-people in the second chamber who can judge when transient impressions occur.

There is another link between Mill and Madison as well. Mill's fear is that absent a second chamber, the first would be despotic. Madison expresses the same concern when he thinks that a second chamber is required to protect the people from those who rule them. We have plenty of reasons to fear the despotism of a primary chamber in the governance of SSI. This can be

[209] The name mixed government is used because the ancient Greek system of governance is still significantly different from contemporary government arrangements, at least in the Western world (Russell, 2001: 19).
[210] Probably the future US legislature, given that the argument was made at the American Constitutional Convention.

seen by the background condition of the perfect moral storm, which creates plenty of incentives for despotism and moral corruption (as was explained in Chapter 3). This is one of the strongest points in favour of a second chamber in the case of SSI: to prevent intergenerational despotism, or what Gardiner calls the 'tyranny of the contemporary'.

To clarify the thought is not that the second chamber would have agents who are better motivated to care about future generations. The thought is that a second chamber can be designed in such a way that the second chamber performs its task of representing the interests of future generations. If this role is performed properly then it should be a counter-weight to the potential despotism of the first chamber. In this sense the chamber is perhaps much more aligned with Madison's vision than Mill's, for Madison's vision does not seem to rely on the second chamber having good states people. Although Mill's idea of good statespeople being in a second chamber has some appeal, and would be welcome if it were to happen, the strength of the second chamber is not contingent upon there being good statespeople in it in this case.

As observed earlier, Aristotle and Mill both offer arguments in favour of a second chamber which stand in stark contrast to each other. Despite the lack of agreement between them, they both identify some desirable features of the second chamber that I am considering in the context of SSI governance. Aristotle is right that it is a particular social group which is going to be represented, namely, future generations.[211] Yet Mill is also right in that a talented states-person may be better placed to do a good job at representing their interests, because of the complexity of these interests (Mill, 1861: chapter 13). Again, there appears to be no reason to think that this argument would only apply to the national level, for nothing about it is contingent on the make-up of national-level governance. There is a specific group – future generations – and talented people may be better placed to represent their interests.

I will now turn to consider the type of concerns that someone could express if they were to use the logic of Abbé Sieyès, one of the political thinkers of the French Revolution. On the subject of second chambers he observed that 'if a second chamber dissents from the first it is mischievous; if it agrees, it is superfluous' (Baldwin, 2001: 171). I think Sieyès's argument does have an intuitive appeal, and so the question to ask is whether it can be articulated in terms of global governance. Certainly, we could argue that if we have a primary decision-making chamber which is sufficiently representative of the present generation and if the will of this

[211] Future generations may not be understood as a social group in the classic sense, but they are the type of group which Aristotle seems to have in mind; namely, a group other than that which is already represented in the decision-making process.

chamber is thwarted by a second chamber then it would seem that the second chamber had created a problem. Likewise, if the second chamber agreed with the first then it would not seem to have contributed anything, so what would be the point of it?

Yet I do not think the argument has much to offer in our particular context.[212] Even if a second chamber were only ever to agree with the decisions of the first, it could offer more legitimacy to that decision by confirming that the decision does not violate the interests of future people and therefore does not violate the basic demands of intergenerational justice. And if the second chamber disagreed with the decision of a first, this does not necessarily make it mischievous. In context, it just means that the first chamber has failed to make a decision which is compatible with the relevant interests of future people.[213] To label this type of second-chamber function as mischievous is to make a strong claim that the present generation owes less to future people than the protection of their interests.

A more contemporary thinker, Meg Russell identifies three key benefits that a second chamber can offer: detailed scrutiny, a different perspective, and asking the primary chamber to think again (Russell, 2000: 262). The first point seems to be in the spirit of Mill and Madison, that of preventing despotism. In order to check whether proposals are despotic or not, the chamber provides a preventative function favoured by the need to scrutinise proposals which come before the chamber. The point about providing a different perspective follows closely on from this scrutiny, since the scrutiny is done with the perspective of the interests of future generations in mind. Russell's third point is premised on a relation between the two chambers where the second has power to send decisions back to the first and ask it to reconsider the whole or parts of the decision. Again this is the type of the power which a second chamber could have if it deems that a decision does violate the interests of future generations as they are set out in the declaration of what is owed to the future. Russell captures three of the key benefits of a second chamber in our case.

To summaries, there are a range of interesting and perhaps conflicting arguments which have been offered to defend the use of second chambers. Madison and Mill offers a particularly compelling argument which is that a second chamber can be designed to prevent the despotism

[212] Although I think it definitely makes sense to apply this argument to chambers at the global level for the reasons given in the earlier paragraph about Madison. However, the relevance of the argument is due to how it offers a commentary on the relationship between two chambers, not because of the chamber being at the national level.
[213] This has parallels with how second chambers can be understood as constitutional guardians. The parallel is that they seek to protect something, as opposed to advocating particular policies.

of the first. This thought is combined with an idea which can be seen in the work of Aristotle, which is that the chamber can represent the interests of a particular group such as future generations. Whilst the arguments from these thinkers seem to be designed with second chambers at the national level in mind,[214] there does not appear to be a reason as to why they can be applied to a second chamber at the international level.[215]

[214] With the possible exceptions of Aristotle who may have had city states in mind as oppose to nations, given the time that he was writing.

[215] It is important to note that this thesis is just focused on SSI governance not broader question of international climate or environmental governance. The arguments about a second chamber presented in this chapter may be applicable to these broader questions. Yet exploring the possibility of applying this second chamber argument to environmental or climate governance is beyond the scope of this thesis. There may even be an oddness to accepting the view that that SSI governance should have a second chamber yet that other significantly future effecting issues at the global should not.

6.6 Would a second chamber also suffer from moral corruption?

One may wonder whether the problem of moral corruption is going to be relevant in the second chamber. The potential for moral corruption clearly exists in this chamber, and is something that a well-designed second chamber would have mechanisms to guard against. Perhaps a variation of the transparency, publicity and accountability solution presented in chapter 3 would be appropriate for guarding against moral corruption in a second chamber as well. Yet we do have reason to think that members of the second chamber would be less susceptible.

One reason for thinking that a second chamber would be less likely to experience moral corruption is that the members would not be subject to as many incentives for short-termism, which invites moral corruption.[216] Short-termism is when net benefits are given to the present generation at the expense of future one. Elections invite short-termism for many reasons, ranging from voter preferences, to political gamesmanship, to the force of special interest groups (Mackinze, 2016: 24-27).[217] Consider the point that it is hard for politicians to persuasively communicate the idea that policy will produce a benefit for the future, this is difficult due to the complicated nature of such claims, and the uncertainty to which they are subject (Mackinze, 2016: 27). The thought is that it is easier for politicians to persuade the electorate of the possibility of net present benefits than net future benefits. This is one amongst many reasons why elections could invite short-termism and moral corruption, for they invite a focus on net-present benefits. A second chamber would not be subject to elections so avoids this source of short-termism, thereby avoiding a source of moral corruption.

A more powerful reason for thinking that moral corruption would not be as prevalent in the second chamber is the remit of this chamber. The second chamber is future generations focused and, more importantly, it is future generations focused with a minimally just Statement of What is Owed to the Future reminding the members of their task, which is to protect the content of that statement. Contrast this to a primary chamber, which is not bound to pay attention to *A Statement of What is Owed to the Future*. The primary chamber has a huge and complicated range of interests to account for, thus it is understandable (although not permissible) for the interests of future generations to be neglected. Such moral corruption may occur through negligence, by failing to give due consideration to future generations amongst all the other

[216] Moral corruption could occur when justification for the short-termism is offered.
[217] For a proper exploration of the sources of short-termism see Mackenize (2016).

interests they consider, or at the other end of the spectrum, moral corruption could be a product of a more sinister disregard of the worth of future generations.

This also shows that the reasons to think that the second chamber is less susceptible to moral corruption are independent of the moral character of the agents; it is not a claim that the second chamber would have more virtuous agents. If you were to have equally virtuous agents we will still have more corruption in the primary chamber, due to the way that the respective chambers are designed.

6.7 A link between these proposals and *A Statement of What is Owed to the Future*

I shall briefly comment on the relationship between my proposals for giving weight to the interests of future people and *A Statement of What is Owed to the Future*. In short, *A Statement of What is Owed to the Future* may provide a suitable account of the interests which could be protected by the proposals for giving weight to the interests of future people.

A Statement of What is Owed to the Future provides a heuristic for thinking about future people. Thinking about the future is challenging, and this should therefore ease the challenge for policy-makers. The statement will provide a clear account of the relevant interests of future people and by doing so it should help policy-makers think about future people in the same way that any clear document about X should help a non-expert to think about X.

Yet the benefit is more than this. *A Statement of What is Owed to the Future* does not only provide a clear way to think about the interests of future people; it could also engage with some interesting psychological mechanisms. First, it could produce an anchoring effect, and it should encourage policy-makers to think about future people in this particular way. This means that when decision-makers are faced with a decision about future people, they should be able to think about it with *A Statement of What is Owed to the Future* in mind, because this will be the first standard against which the policy should be compared. This assumption is based on an understanding of human psychology, which points to the informative power of the first piece of information which is made available to an agent about a particular topic.

There is reason to think that a priming effect will be produced by *A Statement of What is Owed to the Future*. That is to say, the existence of *A Statement of What is Owed to the Future* is going to mean that people are more likely actually to consider what is owed to the future. And this need not be restricted to SSI if it is clear that the statement has applicability to other questions of governance and future people.

There is also reason to think that *A Statement of What is Owed to the Future* will reduce friction around talking about the interests of future people. The idea of friction is anything that presents an obstacle (Leventhal, *et al*, 1965; Thaler and Sunstein, 2008: 75-78). Imagine a situation in

which you wish to have a dialogue about how action X affects the interests of future people. In this situation, one likely obstacle concerns what is meant by 'the interests of future people'. However, if there is already a document in which the interests of future people are clearly stated, then a significant obstacle to the conversation is removed.

This is not the place for a complete psychological analysis of the impact which *A Statement of What is Owed to the Future* may have. Nonetheless, as outlined above, we do have reason to think that such a document could help us engage with those facets of human psychology which would persuade the present generation to take the minimal interests contained within *A Statement of What is Owed to the Future* seriously.

6.8 Can future generations genuinely be included within an SSI governance arrangement?

Are there genuinely satisfactory ways for the inclusion of future generations within an SSI governance arrangement? As with so many issues, the answer to this question is ‘yes’ in some senses and ‘no’ in others.

If we think of the ideal of inclusion, then the answer is ‘no’. As explained at the beginning of this chapter, the ideal of inclusion can be met by the meaningful participation of that group which is meant to be included. Yet in the case of future generations this is clearly impossible, and therefore we must settle for representation as the tool for including future generations. Insofar as representation is not the ideal form of inclusion, then we can express dissatisfaction with representation being the tool which has to be used to try to achieve inclusion.

It is perhaps important be aware of and express the imperfect nature of representation as a way of including the interests of future generations. This may be valuable to avoid complacency in our treatment of these interests. Nonetheless we would be also be in remise if we were to use the imperfect nature of representation to lambast an SSI governance institution which sought to account for interests of future generations in this way. Firstly it would be making a positive move to include the interests of a vulnerable group, and that is a valuable endeavour. Secondly many governance systems use representation as a form of including the interests of people. A critique of representation as form of inclusion should have a much broader target than just an SSI governance institution.

‘

6.9 A criticism: the proliferation of the second chamber as a problem-solving tool – how many chambers do you want?!

I shall now consider some of the criticisms of using a second chamber as a way to address the problem of inclusion.

The first criticism I will consider relates to why a second chamber (rather than some other possible option) should be used to address the problem of the inclusion of future generations. Second chambers are can be powerful bodies, and there is a plethora of problems which they could be used to engage with, such as preventing the weaponising of SSI, or trying to achieve the most just state of affairs possible, or being a more inclusive place of deliberation; a place where every marginalised voice has a chance to speak. The powers of this chamber are absolutely huge. Hence we have to ask whether it is justifiable or defensible to provide such a narrow focus for the functioning of a second chamber.

A criticism with perhaps a similar starting point is that if we allow future generations to have their own chamber, then every group with a cause will want their own chamber, and we will end up with a system in which there are 50-plus decision-making chambers. Such a system seems to be undesirable on practical grounds alone, for a system with 50-plus decision-making chambers would be in danger of being highly slow and inefficient. Therefore one may argue that it is better not to entertain the notion of a second chamber because it is just the start of a slippery slope.

The response to both of these criticisms starts by highlighting the nature of the second chamber which is being proposed. It is a chamber which is meant to protect the basic interests of future generations. It serves to prevent profound injustices to future people. It is not about creating an optimal state of affairs or achieving complete justice for anyone; it is about the protection of a minimal set of interests which ought not to be violated. These interests have a special status in that they are a basic component of justice. The other concerns raised do not have that same status, and therefore they do not have the same need to be protected through a second chamber.[218] Of course, there may well be other considerations of basic justice that may deserve protection through a second chamber, such as marginalised voices and non-human nature. This

[218] Additionally, these other concerns could be subject to different solutions. For example, there may be other tools for bringing marginalised voices into the decision-making process which do not require a second chamber.

would lead to questions of how a second chamber could be designed which protects all components of basic justice. This is not a question I can answer here, although it is one which is clearly important, and the account of how a second chamber could function for protecting the interests of future generations is clearly relevant for such a question, insofar as protecting the interests of future generations is a very important part of basic justice.

6.10 Conclusion

To conclude this chapter, it is clear that the abstract principle of including future generations is one which can be realised in a plethora of ways. I have argued in favour of starting with a simple idea in the form of *A Statement of What is Owed to the Future*. This statement can provide a strong foundation for some emerging tools which can be used to protect people's interests: ombudspersons and constitutional measures. Moreover, this statement could also have other benefits, for example by reducing friction between policy-makers who engage with the interests of future people. It should also have a priming and an anchoring effect, as explained in this chapter.

The chapter has considered how the principle of inclusion should be understood, and it has identified participation as the ideal way of meeting the standard for a principle of inclusion. However, given the impossibility of achieving this, the chapter has focused on representation instead as the way of meeting the standard for a principle of inclusion. In order to do so, the chapter has considered using a second chamber as a way of accounting for the interests of future people in an SSI governance decision-making procedure. The reason for selecting a second chamber as a suitable mechanism is based on the argument that a second chamber is most likely to provide a high quality of deliberative space to consider the interests of future generations, since that would be the sole purpose of the chamber.

Finally it is worth recalling that how this chapter understands the argument which is presented in favour of a second chamber. It has presented a political theory argument in favour of a second chamber. Yet prior to being able to determine if a second chamber is a suitable way of protecting the interests of future generations there would have to be an engagement with the experts on institutional design. Therefore I understand this a chapter as presenting a case for thinking that a second chamber is a good way of representing the interests of future people. If this argument is accepted the next step is to engage with the experts on institutional design.

7 Conclusion

This thesis is motivated by the importance of having a clearer picture of some of the ethical worries about SSI that its governance must address. To contribute to this picture the thesis examined three under explored issues that an ethical SSI governance arrangement should address. Namely those of moral corruption, moral hazard, and the inclusion future generations.

The moral corruption chapter focused on how moral corruption could be addressed. This is useful insofar as addressing moral corruption appears to be absent from governance reports on SSI, and the political theory literature. The argument in this chapter was that a well-functioning accountability mechanism could be well-placed to address moral corruption. The chapter drew attention to the fact that transparency is already a common proposal in governance reports on SSI, so these reports may at first glance may already be well- designed to address moral corruption as well. The chapter argued this was not necessarily the case due to the reports tending to have articulations of transparency which failed to appreciate the relationship that it has with publicity and accountability. Therefore this chapter offers two key points; one that well-functioning accountability is a promising was to address moral corruption, two that well-functioning accountability requires transparency and publicity and that this point is neglected in current proposals for transparency in SSI governance reports.

The first moral hazard chapter focused on how the moral hazard complaint should be understood. The chapter was a product of the short-comings from the governance reports and philosophical literature on SSI and the moral hazard, where a clear and satisfactory account of the hazard appears to be absent. This chapter argued that a complete moral hazard complaint should have an answer to 5 different variables; the stage of SSI, who the relevant agents are, the mechanism by which the hazard occurs, what the impact of the hazard is, and why the hazard is bad. By doing this chapter has sought to provide a degree of conceptual clarity to the hazard which is necessary if the hazard is going to be addressed. Moreover this degree of clarity hopefully sheds light on a concept which has a reputation of being vague and ambiguous, and drawn attention to the sources of such ambiguity (disagreement or vagueness about how the variables are understood). Finally the chapter applied Henry Shue's framework of threshold likelihoods to justify taking the hazard as problem seriously even though there is an absence of empirical evidence about whether it does affect all the relevant agents (in our case the formal role holders of power).

The following chapter considered secrecy as a tool to address the moral hazard problem. The reason for taking secrecy is that secrecy has been mentioned in the moral hazard literature as possible solution (Lin, 2013: 709), yet the possibility and desirability of secrecy in the case of SSI governance have not been well explored.[219] The chapter drew on the work of Pozen and applied it to the Manhattan Project to build a picture of the concept and potential mechanism of secrecy. This enabled the chapter to build a scenario of the secrecy scenario in SSI governance which it argued was fit to address the moral hazard that the thesis is interested in as well as many other potential accounts of the hazard. The rest of this chapter focused on considering reasons we have to be opposed to the secrecy scenario. And concluded that unless the likelihood or badness of the hazard becomes much greater, then the secrecy scenario does not appear to be price worth paying to address the hazard, hence it should be absent form SSI governance. A contribution of this chapter is that it elucidate that idea of secrecy in the context of SSI governance, and may by doing so provide a clearer picture of what it is that many of us have strong reasons to be opposed to.

The final chapter sought to address the intergenerational problem of inclusion. The motive behind this chapter is that future generations are very important yet it is unclear how future generations should be included in SSI governance, this again is a problem which is true of the philosophical literature and governance reports on SSI. Governance reports on SSI are particularly in remise on this issue where they sometimes pay lip service to the importance of future generations but fail to proposal any meaningful mechanisms to account for them. The chapter in argued in favour of *A Statement of What is Owed to the Future* as a tool which can help facilitate the inclusion of future generations. The argument progressed to ultimately endorse a second chamber for future generations in SSI governance. The contribution of this chapter is that it provides tools and ways to think about including future generations which have been absent from literature and governance reports on SSI governance.

The above few paragraphs have presented the content of these chapters in isolation, whilst they do standalone in addressing issues which are of interest ethic concern for SSI governance, they are also deeply related problems. Here I shall provide a few comments on the relationship which exists between these chapters. Perhaps the most obvious relation is how the secrecy chapter defies the proposals from the moral corruption chapter and the chapter of intergenerational conclusion. It is doubtful that the conditions of secrecy could be met if there

[219] Or if they have, it has been done in secret.

was a well-functioning second chamber to represent the interests of future generations, although this is not logically impossible.

The relationship between the moral corruption and moral hazard concerns has already been explored in chapter four. Where the point is made that moral l corruption may be a type of ethical failure which leads the hazard to occur. The point was that moral corruption degrades the quality of public discourse around SSI. And if the quality of discourse if worse than the formal role holders of power may be more susceptible to acting in a way which has an adverse impact on mitigation levels.

It is worth drawing attention to the relationship between the well-functioning accountability mechanism which was proposed in the moral corruption, and the proposed second chamber to address the problem of intergenerational inclusion. The thought is that a second chamber contributes to a well-functioning accountability. Failure by decision makers to abide by the statement of what is owed to the future in their decision would result at least in those decisions being held to account in the second chamber, or depending on the relationship with that chamber and other courts, the chamber taking that to a court to hold the decision to account. These proposals appear to appear to be complementary to each other

The thesis has provide ways to think about the problems of moral corruption, moral hazard and the intergenerational problem of inclusion, which could contribute to how these problems are addressed in the literature on SSI governance and broader philosophical literature on the these topic.

This thesis, as all works do has its limitations. Two clearly come to mind when considering the moral hazard. The first of which is that it does appear to be contestable how seriously the moral hazard should be taken given that we lack empirical evidence about the likelihood of the relevant agents experiencing the hazard.[220] Secondly the thesis was unable to offer a satisfactory way of addressing the moral hazard problem. It was only able to offer secrecy which we have strong reasons to reject. Additionally the accountability mechanism which was proposed to address the problem of moral corruption is unlikely to be fit to fully address the moral corruption, the chapter offers a partial solution to the problem. Lastly it is clearly not a complete project, a complete project provide guidance on all the ethical concerns with SSI

[220] Although it did try to argue that such evidence was not strictly required via Shue's threshold likelihood framework, nonetheless the case for taking the hazard seriously may be more compelling if there was empirical evidence which provided reason to belief that the relevant agent may experience the hazard.

which were raised in the literature review. This clearly not a complete account of the limitations of the thesis, you the reader, probably have your own list. Nonetheless they are ones which stand out to me as the key limitations of this project.

The last limitation pointed to the fact that there is a plethora of concerns about SSI, and this thesis has only been able to consider three. This points to how this type of project could be developed. Simply by expanding the sphere of ethical issues which attention is given to. One could also look at a different sphere of governance and consider how SSI governance at the national level could be fit to address these ethical concerns. For example, one could ask the question of whether a second or additional chamber at the national level holds any promise for addressing the problem of intergenerational inclusion in SSI governance.

Alternatively one could consider a different type of geoengineering technology and explore whether these problems exist, and whether the proposed ways to address them would work for negative emission technologies. Given the inclusion of negative emission technologies in models to achieve 1.5 degrees this arguably a very pressing area of inquiry. Alternatively this project could be expanded in a different way. Which would be to consider the problems of moral corruption, moral hazard and the intergenerational problem of inclusion, and seeing whether the proposed understandings and solutions to these problems transfer to other issues. For example other new future effecting technologies like AI may raise the intergenerational problem of inclusion in governance, would the tools proposed address the intergenerational problem of inclusion in AI governance?

The day may not be too far in the future when SSI is governed. We have choice at the moment what to do in light of the suspicion that that day is not too far away. I think that we have a duty to ensure that those who seek to govern SSI have the tools they need to do so in an ethical way. The part we can play is by clearly understanding the nature of the problem and thinking of tools and strategies which may be fit to address these problems, and being ready to clearly communicate this to those who seek to govern SSI. This thesis has made a contribution by engaging with a few of the issues. When the dawn of SSI governance emerges we as ethicists must be well prepared to communicate the nature of the problem and potential solutions to those who seek to govern SSI. If we are not, our failure may condemn the world to a state of affairs which no one deserves.

8 The Bibliography

Abelkop A, and Carlson J. (2013) Reining in Phaëthon's chariot: principles for the governance of geoengineering. *Transnational Law & Contemporary Problems*. 21, 763–807.

Adams, D. (1995) *The Hitchhikers Guide To The Galaxy: A Trilogy in Five Parts* Portsmouth, William Heinemann.

Agius, E. (2006). A European Ombudsman for the Rights of Future Generations, in Benedek Jávor and Judit Rácz (eds), *Do we Owe them a Future? The Opportunities of a Representation for Future Generations in Europe* Budapest: Védegylet—Protect the Future! pp.16–30.

Archer, D. and Brovkin V. (2008). The millennial atmospheric lifetime of anthropogenic CO2. *Climatic Change*, 90(3), pp.283-297.

Aristotle ([350B.C] 1998), Politics, trans. C. D. C. Reeve, Indianapolis: Hackett Publishing Company.

Armeni, C. (2015). Global experimentalist governance, international law and climate change technologies. *International & Comparative Law Quarterly*. 64,pp. 875–904.

Arrow, K. (1999). Discounting, morality, and gaming. In P. R. Portney and J. P. Weyant (Eds), *Discounting and Intergenerational Equity*, New York, NY: Resources for the Future, pp.13-21.

Baard P, and Wikman-Svahn P, (2016). Do we have residual obligations to Engineer the Climate as a Matter of Justice? In C Preston, (Eds). *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene*. New York: Rowman & Littlefield, pp. 49-62.

Baatz C and Ott K (2016). Why Aggressive mitigation must be Part of Any Pathway to Climate Justice. In C Preston, (Eds). *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene.* New York: Rowman & Littlefield, pp.93-108.

Baatz, C (2016). Can we have it both ways? On potential trade-offs between mitigation and Solar Radiation Management. *Environmental Values* 25 (1) 29-49.

Baer P, Athanasiou T, Kartha S and Kemp-Benedict E, (2010). Greenhouse Development Rights: A Framework for Climate Protection That is "More Fair" Than Equal Per Capita Emissions Rights. In Gardiner, S, Caney, S, Jamieson, D and Shue, H (Eds). *Climate Ethics: Essential Readings*, Oxford: Oxford University Press, pp.215-231.

Baker, T. (1996). On the Genealogy of Moral Hazard. *Texas Law Review*. 75 (2): 237-92.

Baldwin, N. (2001). Concluding Observations. In D Shall and N. D. J. Baldwin (Eds). *Second Chambers*. London: Frank Cass & Co., pp. 171-180.

Barrett, S. (2008). The Incredible Economics of geoengineering. *Environmental and resource economics* 39: 45-54.

Barrett S, Lenton T, and Millner A. (2014). Climate engineering reconsidered. *Nature Climate Change* 4: 527–529.

Barry, B. (1999) Sustainability and intergenerational justice, in A. Dobson (eds.), *Fairness and Futurity*, Oxford: Oxford University Press. pp. 93–117

Barry, B. (2005) *Why Social Justice Matters* London: Polity.

Beckerman, W. (1999). Sustainable development and our obligations to future generations. In A. Dobson (Eds.), *Fairness and Futurity: Essays on Environmental Sustainability and Social Justice*. New York: Oxford University Press. pp. 71–92.

Beckman and Uggla. (2016). An Ombudsman for Future Generations: Legitimate and Effective? In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press, pp. 117-134.

Benedick RE. (2011). Considerations on governance for climate remediation technologies: lessons from the 'ozone hole'. *Stanford Journal of Law, Science and Policy* 4: 6–9.

Bentham, J (1931 [1802]). *The Theory of Legislation*, C.K. Ogden (eds). London: Routledge and Kegan Paul.

Bidadanure, J. (2016). Youth quotas, diversity, and long-termism. In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press pp.266-281.

Biondi, A. C. (2007). Aristotle on the mixed constitution and its relevance for American political thought. *Social Philosophy & Policy Foundation* 24 (2): 176-198.

Bipartisan Policy Center (2011) Geoengineering: A National Strategic Plan for Research on the Potential Effectiveness, Feasibility, and Consequences of Climate Remediation Technologies. Washington, DC.

Bodansky, D, (1996), May we engineer the climate? *Climatic Change*, 33: 309-321.

Bodansky, D. (1999). The Legitimacy of International Governance: A Coming Challenge for International Environmental Law? *The American Journal of International Law* 93, (3): 596–624.

Bodansky, D. (2013). The who, what, and wherefore of geoengineering governance. *Climatic Change*, 121(3): 539-551.

Bodle, R, ,Oberthür S, Donat L, Homann G, Sina S, and Tedsen E. (2014) *Options and Proposals for the International Governance of Geoengineering*. Berlin: Ecologic Institute

Boissay, F, Collard F and Smets F. (2016). Booms and banking crises. *Journal of Political Economy*, 124: 489– 538.

Bok, S, (1982). Secrets: *On the Ethics of Concealment and Revelation*. New York: Pantheon books.

Bracmort, Kelsi, R. K. Lattanzio, and E. C. Barbour (2010) *Geoengineering: Governance and Technology Policy*. Washington, DC: Congressional Research Service.

Broome, J, (1992). *Counting the Cost of Global Warming*. Isle of Harris, White Horse Press.

Broome, J. (1994). Discounting the future. *Philosophy & Public Affairs* 23(2): 128-156.

Broome, J. (2012*) Climate Matters: Ethics in a Warming World* New York: Norton & Company

Brovkin, V. Petoukhov V, Claussen M, Bauer E, Archer, D, and Jaeger, C. (2009). Geoengineering climate by stratospheric sulfur injections: Earth system vulnerability to technological failure. *Climatic Change*, 92: 243–259.

Brown, Campbell, (2005), Priority or Sufficiency ...or Both?, *Economics and Philosophy*, 21: 199–220.

Buchanan, A, and Keohane R, O. (2006). The Legitimacy of Global Governance Institutions. *Ethics and International Affairs.* 20 (4): 405–37.

Buck, HJ (2012) Climate remediation to address social development challenges: Going beyond cost-benefit and risk approaches to assessing solar radiation management. In C Preston, (Eds). *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene*. New York: Rowman & Littlefield pp. 133–148.

Bunzl M. (2009) Researching geoengineering: should not or could not? *Environmental. Research Letters,* 4, (4).

Cairns R, and Nightingale P.(2014). *The security implications of geoengineering: Blame, imposed agreement and the security of critical infrastructure*. (Climate Geoengineering Governance Working Paper 18). See http://www.geoengineering-governance-research. org/perch/resources/workingpaper18nightingalecairnssecurityimplications.pdf (accessed 19 December 2019).

Callies, D,.(2019a). *Climate Engineering: a Normative Perspective*, London, Lexington Books.

Callies, D. (2019b). The Slippery Slope Argument against Geoengineering Research. *Journal Applied Philosophy*, 36: 675-687.

Caney, S. (2010). Climate Change, Human Rights, and Moral Thresholds. In Gardiner, S, Caney, S, Jamieson, D and Shue, H (Eds). *Climate Ethics: Essential Readings*, Oxford: Oxford University Press, 2010 pp.163-180.

Caney, S. (2012). Just Emissions. *Philosophy & Public Affairs*, 40(4), 255-300.

Caney, S. (2014). Climate change, intergenerational equity and the social discount rate. *Politics, Philosophy & Economics,* 13(4): 320-342.

Caney, S. (2016). Political Institutions for the Future: A Fivefold Package. In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press, pp. 135- 155.

Caney, S. (2018) Global Governance: Procedures, Outcomes and Justice Cabrera, L Cabrera, (Eds), *Institutional Cosmopolitanism,* Oxford University Press, Oxford.

Carbon Brief (2014) *Two degrees: The history of climate change's speed limit*. [online], UK available at https://www.carbonbrief.org/two-degrees-the-history-of-climate-changes-speed-limit [accessed on 20 /11/2016].

Carr, Wylie A. and Preston, C. (2017). Skewed Vulnerabilities and Moral Corruption in Global Perspectives on Climate Engineering. *Environmental Values* 26 (6): 757–77.

Casal, Paula, (2007), Why Sufficiency Is Not Enough, *Ethics*, 117: 296–326.

Chhetri N, Chong D, Conca K, Falk R, Gillespie A, Gupta A, Jinnah S, Kashwan P, Lahsen M, Light A, McKinnon C, Pauliele L, Valdivia W, Wapner P, Morrow D, Turkaly C, and Nicholson S. (2018). *Governing Solar Radiation Management. Forum for Climate Engineering Assessment*, American University. October. http://ceassessment.org/ SRMreport.

Cohen, G. (2008). *Rescuing Justice and Equality*. Cambridge, Mass.: Harvard University Press.

Corner, A. and Pidgeon N. (2014). 'Geoengineering, climate change scepticism and the "moral hazard" argument: an experimental study of UK public perceptions'. *Philosophical Transactions of the Royal Society* A 372 (2031).

Cotterrell, R. (2000) Transparency, mass media, ideology and community. *Cultural Values* 3: 414–426.

Cressy, D. (2012). Geoengineering experiment cancelled amid patent row. *Nature,* 15 May.

Cripps, E. (2013) *Climate change and the Moral Agent: Individual Duties in an Independent World.* Oxford, Oxford University Press.

Cripps, E. (2016) On climate matters: offsetting, population, and justice. *Midwest Studies in Philosophy* 40: 114–128.

Crutzen, P. (2006) Albedo enhancement by stratospheric sulphur injections: A contribution to resolve a policy dilemma? *Climate Change, 77*(3-4): 211-220.

Dasgupta, P. (2012). Time and the generations. In R. Hahn and A. Ulph (Eds), *Climate Change and Common Sense: Essays in Honor of Tom Schelling*, Oxford: Oxford University Press. pp. 101-130.

Davis, G. (2013). The Psychological Costs of Geoengineering: Why It may Be Hard to Accept Even If It Works. In W. C, G, Burns and A, L, Strauss (eds). *Climate Change Geoengineering: Philosophical Perspectives, Legal Issues, and Governance Frameworks*. New York: Cambridge University Press, pp.59-77.

De Lazari-Radek, K and Singer, P. (2010), Secrecy in Consequentialism: A defence of Esoteric Morality, *Ratio,* 23: 34-58.

De Lazari-Radek, K and Singer, P, (2014), *The Point of View of the Universe: Sidgwick and Contemporary Ethics* (Oxford, Oxford University Press).

Deweese-Boyd, Ian, "Self-Deception", *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/fall2017/entries/self-deception/

Dobson, A. (1996). Representative democracy and the environment. In W. M. Lafferty & J. Meadowcroft (Eds), *Democracy and the environment: Problems and prospects*. Cheltenham, England: Edward Elgar. pp. 124–139.

Dubner, S.J. and Levitt, S.D. (2009). *SuperFreakonomics. Global Cooling, Patriotic Prostitutes and why Suicide Bombers Should buy Life Insurance.* HarperCollins Publishers, Toronto.

Duff, Antony. (2007). *Answering for Crime: Responsibility and Liability in the Criminal Law*. Oxford: Hart Publishing.

Dunlap RE, McCright AM (2011) Organized climate change denial. In: Dryzek J, Norgaard R, Schlosberg D (Eds) *The Oxford Handbook of Climate Change and Society*. Oxford University Press, Oxford, pp. 144–160.

Dunlap RE, Jacques PJ (2013) Climate change denial books and conservative think tanks: exploring the connection. *American Behavioral Scientist*, 57(6): 699–731.

Ekeli, K. (2016). Electoral Design, Sub-Majority Rules, and Representation for Future Generations. In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press, pp. 214- 227.

Elliot K. 2010 Geoengineering and the precautionary principle. *International Journal of Applied Philosophy*. 24: 237–253.

Etzioni, A. (2010) Is transparency the best disinfectant? *Journal of Political Philosophy* 18: 389–404.

Fearon, J, (1998) Bargaining, Enforcement, and International Cooperation, *International Organization* 52(2): 269–306.

Feinberg, J. (1970*). Doing and deserving: Essays on a Theory of Responsibility* (Princeton, Princeton University Press).

Fioretos, O, (2011) "Historical Institutionalism in International Relations." *International Organization* 65(2): 367-99.

Fleming, J. (2010) *Fixing the Sky: The Checkered History of Weather and climate control*, New York: Columbia University Press.

Fox, J (2007) The uncertain relationship between transparency and accountability. *Development in Practice* 17: 663–671.

Fragnière, A and Gardiner S. (2016). Why Geoengineering is not "Plan B". In C Preston (Eds) *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene*. New York: Rowman & Littlefield International Ltd, pp. 15-30.

Freedman, R. (2014). *Failing To Protect: The UN and the Politicisation of Human Rights*. London: Hurst and Co.

Gardiner S, Caney S, Jamieson,D. and Shue H. (2010). *Climate Ethics*. Oxford: Oxford University Press.

Gardiner, S. (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change.* Oxford: Oxford University Press.

Gardiner, S. (2013). Geoengineering and Moral Schizophrenia: what is the Question? In W. C, G, Burns and A, L, Strauss (eds). *Climate Change Geoengineering: Philosophical Perspectives, Legal Issues, and Governance Frameworks*. New York: Cambridge University Press, pp. 11-38.

Gardiner, S. (2014). 'A Call for a Global Constitutional Convention Focused on Future Generations'. *Ethics & International Affairs*, 28(3): 299-315.

Gardiner, S, (2016), *In Defense of Climate Ethics,* In Gardiner, S.,& Weisbach, D. A. (2016). *Debating Climate Ethics*. New York: Oxford University Press.

Gardiner, S, Fragnière A. 2018 The Tollgate Principles for the governance of geoengineering: moving beyond the Oxford Principles to an ethically more robust approach. *Ethics, Policy and Environment*. 21: 143–174.

Geoengineering for Decision-Makers: Science and Technology (2011)

Gidley J.M, Fien J, Smith, J.-A, Thomsen D.C, and Smith T.F. (2009), Participatory futures methods: towards adaptability and resilience in climate-vulnerable communities. *Journal of Environmental Policy and Planning*, 19: 427-440.

Gifford, R. (2011). The dragons of inaction: psychological barriers that limit climate change mitigation. *American Psychologist*. 66:290–302.

Gilabert P, and Lawford-Smith, H, (2012). Political Feasibility: A Conceptual Exploration, *Political Studies*, 60: 809–25.

Gingrich, N (2008) "Stop the Green Pig: Defeat the Boxer-Warner-Lieberman Green Pork Bill Capping American Jobs and Trading America's Future", Human events, June 3, 2008 https://humanevents.com/2008/06/03/stop-the-green-pig-defeat-the-boxerwarnerlieberman-green-pork-bill-capping-american-jobs-and-trading-americas-future/ [accessed on the 16th of May 2018].

Goeschl T, Heyen D, Moreno-Cruz J. (2013). The intergenerational transfer of solar radiation management capabilities and atmospheric carbon stocks. *Environmental and Resource Economics.* 56: 85–104.

Gonzalez-Ricoy, I. (2016). Constitutionalizing Intergenerational Provisions. In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press pp. 170-183.

Gonzalez-Ricoy, I. and Gosseries, A. (Eds) (2016). *Institutions for Future Generations*. Oxford: Oxford University Press.

Gonzalez-Ricoy I, and Rey F (2019). Enfranchising the future: Climate justice and the representation of future generations. *WIREs Climate Change*.;10:e598

Göpel, Maja. (2012). *Ombudspersons for Future Generations as Sustainability Implementation Units.* London: Stakeholder Forum.

Gosseries, A. (2008). Constitutions and Future Generations. *The Good Society* 17: 32–7.

Gosseries, A. (2014) The Intergenerational Case for Constitutional Rigidity. *Ratio Juris,* 27(4): 528–539.

Grant, R. W., and Keohane, R. O. (2005). "Accountability and Abuses of Power in World Politics." *American Political Science Review* 99(1): 29–43.

Grey, C. (2014) An organizational Culture of Secrecy: The Case of Bletchley Park. *Management & Organizational History,* 9:107:122.

Griffin, R. C. (1998), The fundamental principles of cost-benefit analysis, *Water Resources Research.,* 34(8): 2063– 2071.

Grimmelikhuijsen, SG, Welch, EW (2012) Developing and testing a theoretical framework for computer-mediated transparency of local governments. *Public Administration Review* 72: 562–571.

Gutmann A, and Thompson D. (1996). *Democracy and Disagreement: Why Moral Conflict Cannot Be Avoided in Politics, and What Should Be Done about It*. Massachusetts: Harvard University Press.

Gutwald, R, O Leßmann, T Masson, and F Rauschmayer. (2014) A Capability Approach to Intergenerational Justice? Examining the Potential of Amartya Sen's Ethics with Regard to Intergenerational. *Issues, Journal of Human Development and Capabilities*, 15(4), 355-368.

Hale, B. (2012). The World That Would Have Been: Moral Hazard Arguments Against Geoengineering. In C Preston (Eds). *Engineering the Climate: The Ethics of Solar Radiation Management*. Plymouth: Lexington Books, pp. 113-132.

Hale, T. (2008) Transparency, accountability, and global governance. *Global Governance: A Review of Multilateralism and International Organizations* 14: 73–94.

Hall, C. (1998). Institutional Solutions for Governing the Global Commons: Design Factors and Effectiveness. *The Journal of Environment & Development, 7*(2): 86–114.

Hamilton, C. (2013a), *Earthmasters: The Dawn of the Age of Climate Engineering*, New Haven: Yale University Press.

Hamilton, C. (2013b). The Ethical Foundations of Climate Engineering. In W. C. G. Burns and A L. Strauss (eds). *Climate Change Geoengineering: Philosophical Perspectives, Legal Issues, and Governance Frameworks*. New York: Cambridge University Press, pp. 39-58.

Hardin, G. (1968). The Tragedy of the Commons. *Science* 162: pp. 1243–1248.

Hart, H. (2008), *Punishment and Responsibility: Essays in the philosophy of law*, Oxford, Oxford University Press.

Hartman, L. (2017). Climate engineering and the playing god critique. *Ethics & International Affairs*, 31 (3), 313–333.

Hatfield J, and Job R. (2001). Optimism bias about environmental degradation: the role of the range of impact on precautions. *Journal of Environmental Psychology* 21: 17–30.

Hartzell-Nichols, L. (2013),'the' precautionary principle to precautionary principles. *Ethics Policy and Environment*. 16, 308–320.

Heald, D. (2006). Varieties of Transparency In Hood, C and Heald, D (Eds) *Transparency: The Key to Better Governance? Proceedings of the British Academy 135*, Oxford: Oxford University Press. 25–43.

Held, PJ. McCormick F, Ojo A ,and Roberts JP (2016). A Cost-Benefit Analysis of Government Compensation of Kidney Donors. *American Journal of Transplantation*; 16:877– 885.

Heyward, C. (2013). Situating and abandoning geoengineering: a typology of five responses to dangerous climate change. *Political Science & Politics*, 46: 23– 27.

Heyward, C. (2015); Is there Anything New Under The Sun? Exceptionalism, Novelty, and Debating Geoengineering Governance, In Maltais, A and Mckinnon, C (Eds) *The Ethics of Climate Governance,* Lanham, Rowan Littlefield Publishers, pp. 135-154.

Heyward, C. (2018) *Sustainable Development Goals and Geoengineering ,* Guest speaker at the Climate Justice conference between the Universities of Graz and Reading.

Heyward C, and Roser D. (Eds.). (2016).*Climate justice in a non-ideal world*. Oxford, Oxford University Press.

Heyward C, and Rayner S. (2016). Uneasy Expertise: Geoengineering, Social Science, and Democracy in the Anthropocene. In Michael Heazle and John Kane, *Policy Legitimacy, Science and Political Authority: Knowledge and Action in Liberal Democracies* London: Routledge pp.101–121.

Hood, C. (2006). Transparency in Historical Perspective. In Hood, C and Heald, D (Eds) *Transparency: The Key to Better Governance? Proceedings of the British Academy 135*, Edited by: Hood, C. and Heald, D. Oxford: Oxford University Press. pp3–24.

Hood, C. (2010) Accountability and transparency: Siamese twins, matching parts, awkward couple? *West European Politics* 33: 989–1009.

Horizon. (2016) *The Wildest Weather in the Universe*, 2016. BBC 2 Television, 23rd October 2016.

Horton, J and David K. (2016). Solar Geoengineering and Obligations to the Global Poor. In C, Preston, (Eds). *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene.* New York: Rowman & Littlefield, pp79-92.

Horton J, Reynolds J, Buck H, Callies D, Schaefer S, Keith D, and Rayner, S. (2018). Solar geoengineering and democracy. *Global Environmental Politics*, 18: 5–24.

Hourdequin, M. (2012). Geoengineering, Solidarity and Moral Risk. In C Preston (Eds). *Engineering the Climate: The Ethics of Solar Radiation Management*. Plymouth: Lexington Books, pp.15-32.

Hourdequin, M. (2016). Justice, Recognition, and Climate Change. *C Preston*, (*Eds)*. *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene.* New York: Rowman & Littlefield pp.33-48.

Hourdequin, M. (2018). Climate change, Climate Engineering, and the 'Global Poor': What Does Justice Require? *Ethics, Policy & Environment*, 21(3): 270–288

Hourdequin, M. (2019). Geoengineering Justice: the Role of Recognition. *Science Technology Human Values* 44:448–77.

Hulme, M. (2014) *Can Science Fix Climate Change? A Case Against Climate Engineering.* Cambridge: Polity Press.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (Eds.)].

Ipsos-MORI (2010) Experiment Earth: Report on a public dialogue on geoengineering. Natural Environment Research Council, Swindon

Irvine P. J, Sriver R. L, and Keller K. (2012). Tension between reducing sea-level rise and global warming through solar-radiation management. *Nature Climate Change*, 2, 97–100

Jamieson, D (1990), Managing the Future: public policy, scientific uncertainty and global warming. In D Scherer (Eds) *Upstream/downstream: Essay in Environmental Ethics* (Philadelphia, temple University Press).

Jamieson, D. (1991), The Epistemology of Climate Change: Some Morals for Managers, *Society and Natural Resources*; 4:319-329.

Jamieson, D. (1992), Ethics, Public Policy, and Global Warming, *Science Technology and Human Values*, 17: 139-153.

Jamieson, D, (1996). Intentional Climate Change, *Climatic Change*, 33: 326–36.

Jamieson, D. (1998), Global Responsibilities: Ethics, Public Health and Global Environmental Change .*Indiana Journal of Global Legal Studies*, 5:99-199.

Jamieson, D. (2014). *Reason in a Dark Time*. Oxford: Oxford University Press.

Jávor, B. (2006). Institutional Protection of Succeeding Generations: Ombudsman for Future Generations in Hungary, in Joerg Chet Tremmel (Eds.), *Handbook of Intergenerational Justice* Cheltenham: Edward Elgar Publishing, pp. 282–98.

Jefferson T. 1999 (1789). To James Madison, Paris, September 6, 1789.In J Appleby, T Ball (Eds), *In Jefferson: Political Writings*. New York: Cambridge University. Press .pp. 593–98.

Jones A, Haywood J. M, Alterskjær K, Boucher O, Cole J. N. S, Curry C. L. *et al*. (2013). The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP). *Journal of Geophysical Research: Atmospheres*, 118(17), 9743–9752.

Jones, J. V.C.(1985), *United States Army in World War II: Special Studies—Manhattan: The Army and the Atomic Bomb* Center of Military History, United States Army, Washington,

Jordan A, Huitema D, Van Asselt H, and Forster J.(Eds). (2018*). Governing Climate Change: Polycentricity in Action?* Cambridge: Cambridge University Press.

Kahan D.M, Jenkins-Smith H, Tarantola T, Silva C.L, and Braman D. (2015). Geoengineering and climate change polarization: Testing a two-channel model of science communication. *Annals of American Academy of Political & Social Science*, 658**:** 192– 222.

Kahneman, D. (2003) Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review* 93 (5): 1449– 1475.

Kant, I. ([1795] 1991) *Perpetual Peace,* in *Kant's Political Writings*. Cambridge: Cambridge University Press.

Kapembwa, J and Wells, J .(2016), Climate Justice for Wildlife: A Rights-Based Account. In Gabreil Garmendia da Trindade & Andrew Woodhall (Eds), *Intervention or Protest: Acting for nonhuman Animals.* pp 359-390.

Karenin, K. (2016). 'Can we Represent Future Generations?' In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press, pp. 83-97.

Keith, D. W. (2000), Geoengineering the climate: History and prospect, Annual Review of Environment and Resources, 25: 245– 284.

Keith, D. (2013). *A case For Climate Engineering*. Cambridge: Massachusetts Institute of Technology University Press.

Keith, D. (2013). *A Case for Climate Engineering*. Cambridge, MA: MIT Press.

Keith, D. (2017) Toward a Responsible Solar Geoengineering Research Program. *Issues in Science and Technology*. 33 (3):71-77.

Keith, D, and MacMartin D. (2015) A temporary, moderate and responsive scenario for solar geoengineering. *Nature Climate Change* 5(3): 201–206.

Keith D, Parson E, and Morgan G. (2010) Research on global sun block needed now, *Nature* 463: 426–7.

Keohane, R. (1984) *After Hegemony: Cooperation and Discord in the World Political Economy* (Princeton: Princeton University Press).

Keohane, R. (1988). International Institutions: Two Approaches, *International Studies Quarterly* 32: 379-96.

Koremenos B, Lipson C, and S Duncan. (2001). The Rational Design of Institutions. *International Organization* 55 (4):761–99.

Koppell, JG. (2005) Pathologies of accountability: ICANN and the challenge of "multiple accountabilities disorder". *Public Administration Review* 65: 94–108.

Kotowitz, Y. (1989) Moral Hazard. In: Eatwell J., Milgate M., Newman P. (Eds) *Allocation, Information and Markets*. The New Palgrave. Palgrave Macmillan, London. pp.207-213.

Kravitz B, Robock A, Oman L, Stenchikov G, and Marquardt A. B. (2009), Sulfuric Acid Deposition from Stratospheric Geoengineering with Sulfate Aerosols, *Journal of Geophysical Research.*, 114, D14109,

Lawford-Smith, H. (2016). Difference Making and Individuals' Climate Related Obligations. In Heyward, C., & Roser, D. (Eds.). *Climate justice in a non-ideal world*. Oxford: Oxford University Press. pp.64-82.

Lawrence, P. (2014).Justice for Future Generations: *Climate Change and International Law*. Cheltenham, England: Edward Elgar.

Lawrence P, and Kohler L. (2018). Representation of Future Generations Through International Climate Litigation: A Normative Framework. *German Yearbook of International Law*, 60: 639–666.

Leiserowitz A, Maibach E, Roser-Renouf C, Smith N. (2010). *Climate Change in the American Mind: Americans' Global Warming Beliefs and Attitudes in June 2010*.Yale University and George Mason University. NewHaven, CT: Yale Project on Climate Change Communication.

Leventhal H, Singer R, and Jones S. (1965). Effects of fear and specificity of recommendation upon attitudes and behavior. *Journal of Personality and Social Psychology, 2*(1): 20–29

Lin, A. (2009). Geoengineering Governance. *Issues in Legal Scholarship*, 8(1).

Lin, A. (2013) Does Geoengineering Present a Moral Hazard*? Ecology Law Quarterly*, 40 (3): 673-712.

Lindstedt C, and Naurin, D. (2010). Transparency Is Not Enough: Making Transparency Effective in Reducing Corruption. *International Political Science Review* 31: 301–322.

Llanillo P, Jones P. D, and Von Glasow R. (2010). The influence of stratospheric sulphate aerosol deployment on the surface air temperature and the risk of an abrupt global warming. *Atmosphere*, 1(1): 62–84.

Lloyd, I. Oppenheimer M. (2014). On the Design of an International Governance Framework for Geoengineering. *Global Environmental Politics*. 14(2): 45–63.

Lo, Y. T. E, Charlton-Perez A. J, Lott F. C, and Highwood, E. J. (2016). Detecting sulphate aerosol geoengineering with different methods. *Scientific reports*, 6: 39169.

Lo, Y. T. E, Charlton-Perez A. J, Highwood E. J, and Lott F. C. (2018). Best scale for detecting the effects of stratospheric sulfate aerosol geoengineering on surface temperature. *Earth's Future*, 6: 1660– 1671

Long, J, and Scott D. (2013) Vested Interests and Geoengineering Research. *Issues in Science and Technology*. 29: 45–52

Mabillard V, and Zumofen R. (2017). The Complex Relationship Between Transparency and Accountability: A Synthesis and Contribution to Existing Frameworks. *Public Policy and Administration*, 32: 110-129.

MacKenzie M. (2016). 'Institutional Design and Sources of Short-Termism?' In Gonzalez-Ricoy, I and Gosseries, A (Eds), *Institutions for Future Generations.* Oxford: Oxford University Press, pp. 24-45.

MacMartin, D.G Wang W, Kravitz B, Tilmes S, Richter J, and Mills M. J (2019). Timescale for detecting the climate response to stratospheric aerosol geoengineering. *Journal of Geophysical Research Atmospheres*, *124*(3): 1233–1247.

MacMartin D. G,Ricke K L and Keith D W. (2018). Solar geoengineering as part of an overall strategy for meeting the 1.5°C Paris target 376 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*

Marchetti, C. (1977). On geoengineering and the CO2 problem. *Climatic Change*, 1(1): pp.59-68.

Matthews H. D, and Caldeira, K. (2007). Transient climate-carbon simulations of planetary geoengineering. *Proceedings of the National Academy of Sciences of the United States of America*, 104(24), 9949–9954.

McCormack C. G, Born W, Irvine P. J, Achterberg E. P, Amano T, Ardron J, *et al*. (2016). Key impacts of climate engineering on biodiversity and ecosystems, with priorities for future research. *Journal of Integrative Environmental Sciences*, 13:2-4: 103–128.

McCusker K. E, Armour K. C, Bitz C. M, and Battisti D. S. (2014). Rapid and extensive warming following cessation of solar radiation management. *Environmental Research Letters*, 9(2): 24005.

McCusker K. E, Battisti D. S, and Bitz C. M. (2012). The climate response to stratospheric sulfate injections and implications for addressing climate emergencies. *Journal of Climate*, 25(9): 3096–3116

McKinnon, C. (2011). Climate change justice: Getting motivated in the last chance saloon. *Critical Review of International Social and Political Philosophy*, 14(2): 195–213.

McKinnon, C (2012), *Climate Change and Future Justice: Precaution, Compensation, and Triage* Oxon: Routledge, 2012

McKinnon, C, (2019). Sleepwalking into lock-in? Avoiding wrongs to future people in the governance of solar radiation management research, *Environmental Politics*, 28 (3): 441-459.

Mclaren, D, (2016). Framing Out Justice: The Post-politics of Climate Engineering discourses. In C Preston, (Eds). *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene.* New York: Rowman & Littlefield, 139-160.

Meijer, A, (2014) Transparency. In: Bovens, M, Goodin, RE, Schillemans, T (Eds) *The Oxford Handbook of Public Accountability*, Oxford: Oxford University Press, pp. 507–524.

Mercer A. M, Keith D.W, and Sharp J.D (2011), Public understanding of solar radiation management, *Environmental Research Letters*, 6 (4).

Merk C, Pönitzsch G, and Rehdanz K, (2016), Knowledge about aerosol injection does not reduce individual mitigation efforts, *Environmental Research Letters* 11(5).

Michener, G. (2011,) FOI laws around the world. *Journal of Democracy* 22: 145–159.

Mill, J.S. ([1861] 1968). *Considerations on representative government*, (eds) Williams, Geraint. London: Dent.

Mirrlees J. A. (1975): The Theory of Moral Hazard and Unobservable Behavior—Part I, *The Review of Economic Studies 66(01)* pp3-12.

Moe, T. (2006) Power and political institutions. In: Shapiro I , Skowronek S and Galvin D (Eds) *Rethinking Political Institutions: The Art of the State*. New York: New York University Press, pp32-71.

Moellendof, D. (2014), *The Moral Challenge of Dangerous Climate Change: Values, Poverty, and Policy* Cambridge: Cambridge University Press.

Morrow, D. Kopp, R, and Oppenheimer, M. (2013). Political Legitimacy in Decisions about Experiments in Solar Radiation Management. In W. C, G, Burns and A, L, Strauss (Eds). *Climate Change Geoengineering: Philosophical Perspectives, Legal Issues, and Governance Frameworks*. New York: Cambridge University Press, pp 146-167.

Morrow, D. (2014) Ethical aspects of the mitigation obstruction argument against climate engineering research. *Philosophical Transactions of the Royal Society. V* 372, 2014.

Morrow, D. (2017). *International governance of climate engineering* (Forum for Climate Engineering Assessment Working Paper Series #1).

Nalam A, Bala G, and Modak A. (2018) Effects of Arctic geoengineering on precipitation in the tropical monsoon regions, *Climate Dynamics* 50: 3375.

National Research Councils Committee on Geoengineering Climate. (2015) Climate intervention: reflecting sunlight to cool Earth. Washington, DC: National Academies Press.

Naurin, D. (2006). Transparency, publicity, accountability—The missing links, *Swiss Political Science Review*, 12 (3): 90- 98.

Necheles E, Burns L, Chang A and Keith D. (2018). Funding for Solar geoengineering from 2008 to 2018. See: https://geoengineering.environment.harvard.edu/blog/funding-solar-geoengineering (Accessed 27 April 2019)

Nicholls, N. (1999). Cognitive illusions, heuristics, and climate prediction. Bulletin of the *American Meteorological Society* 80: 1385–1397.

Nordhaus, W. (2015). *The climate casino*. Connecticut: Yale University Press.

Norris, R. (2004) General Leslie R. Groves and the Scientists, In Cynthia C (Eds). *Remembering the Manhattan Project: Perspectives on the Making of the Atomic Bomb and Its Legacy*, World Scientific Hackensack.

O'Neill, O. (2006) Transparency and the ethics of communication. In: Hood C, Heald D, (Eds) *Transparency: The Key to Better Governance?* Oxford: Oxford University Press, pp. 75–90.

Organisation for the Prohibition of Chemical Weapons, (2018): *Conference State Parties*. [Online] available at https://www.opcw.org/about-us/conference-states-parties .[Accessed 22 April 2018].

Ott, K. (2012). Might Solar Radiation Management Constitute a Dilemma? In C Preston (eds). *Engineering the Climate: The Ethics of Solar Radiation Management*. Plymouth: Lexington Books, pp. 33-42.

Page, E. (2007). *Climate Change, Justice and Future Generations*. Cheltenham: Edward Elgar Publishing.

Page, T. (1983). Intergenerational justice as opportunity, in D. MacLean & P. G. Brown (Eds.), *Energy and the Future* pp. 38–58 (Totowa, NJ: Rowman & Littlefield), pp. 38–58.

Parfit, D. (1984). *Reasons and Person.* Oxford, Oxford University Press.

Parfit D, (2010). Energy Policy and the Further Future: the Identity Problem. In Gardiner, S, Caney, S, Jamieson, D and Shue, H (eds). *Climate Ethics: Essential Readings*, Oxford: Oxford University Press. pp. 112-121.

Parker, A. and Irvine P.J, (2018). The Risk of Termination Shock From Solar Geoengineering. *Earth's Future 6*: 456-467.

Parson, E. (2014) Climate Engineering in Global Climate Governance: Implications for Participation and Linkage. *Transnational Environmental. Law* 3, (1): 89–110.

Pasztor, J. (2017). The Need for Governance of Climate Geoengineering. *Ethics & International Affairs*, 31(4): 419-430.

Pasztor, J. (2019). *What's in a name? Why we became C2G,* {online] C2G website, last updated 10 of June. Available at https://www.c2g2.net/whats-in-a-name-why-we-became-c2g/ [accessed 18th of November 2019]

Ritchie, H. (2017), *How much will it Cost to Mitigating Climate Change*? [online], Our World in Data .Last update in May 2017. Available at https://ourworldindata.org/how-much-will-it-cost-to-mitigate-climate-change . [accessed on the 22nd of June 2018].

Robock, A.(2016). Albedo Enhancement by Stratospheric Sulfur Injections: More Research Needed. *Earth's Future* 4 (12): 644–48.

Robock A, Oman L, Stenchikov GL. (2008) Regional climate responses to geoengineering with tropical and Arctic SO_2 injections. *Journal of Geophysical Research* 113, D16101.

Rogelj J, den Elzen M, Hohne N, Fransen T, Fekete H, Winkler H, Schaeffer R, Sha F, Riahi K, and Meinhausen M. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature.* 534: 631-639.

Roser, D. (2009). The discount rate: a small number with a big impact. In Applied Ethics: Life, Environment and Society, ed. *Center for Applied Ethics and Philosophy*, 12–27. Sapporo: The Center for Applied Ethics and Philosophy, Hokkaido University.

Roser D, and Seidel, C. (2017).*Climate justice: An introduction*. New York, NY: Routledge.

Rousseau, J.J. ([1772] 1985). *The Government of Poland*, trans. W. Kendall. Indianapolis: Hackett.

Russell, M. (2000*). Reforming the House of Lords: Lessons from Overseas*. Oxford: Oxford University Press.

Sagar, R. (2013).*Secrets and Leaks: The Dilemma of State Secrecy*. Princeton: Princeton University Press.

Schäfer S, Lawrence M, Stelzer H, Born W, Low S, *et al*. (2015) *The European Transdisciplinary Assessment of Climate Engineering* (EuTRACE): Removing Greenhouse Gases from the Atmosphere and Reflecting Sunlight Away from Earth.

Scheffler, S. (2018). *Why Worry About Future Generations*. Oxford: Oxford University Press.

Schuppert F, and Stelzer H. (2016) How much risk ought we to take? Exploring the possibilities of risk-sensitive consequentialism in the context of climate engineering. *Environmental Values* 25, 69–90.

Scott, D. (2012). Insurance Policy or Technological Fix: The Ethical Implications of Framing Solar Radiation Management. In C Preston (Eds) *Engineering the Climate: The Ethics of Solar Radiation Managemen*t. Lanham, MD: Lexington Books. pp151-68.

Sen, A.(2006). What Do We Want from a Theory of Justice. *Journal of Philosophy*, CIII (5): 215-38.

Sharot, T. (2011). *The Optimism Bias: A Tour of the Irrationally Positive Brain*, New York: Vintage Books.

Sharot T, Riccardi A.M, Raio C.M, and Phelps, E.A. (2007). Neural mechanisms mediating optimism bias. *Nature*, 450: 102–105.

Shell, D. (2001) The History of Bicameralism. In Shall, D and Baldwin, N (Eds), *Second Chambers*. London, Frank Cass & Co. 5-18.

Shepherd J, Caldeira K, Cox P, Haigh J, Keith D, Launder B, Mace G, *et al*. (2009) *Geoengineering the Climate: Science, Governance and Uncertainty*. London: Royal Society.

Shoham, Shlomo, and Nira Lamay. (2006). Commission for Future Generations in the Knesset: Lessons Learnt, in Joerg Chet Tremmel (Ed.), *Handbook of Intergenerational Justice* . Cheltenham: Edward Elgar, pp. 244–81.

Shue, H. (1980*). Basic Rights: Subsistence, Affluence and U.S Foreign Policy*. Princeton: Princeton University Press.

Shue, H. (1992). The unavoidability of justice. *The International Politics of the Environment*, 373, pp 373-397.

Shue, H. (1993), Subsistence Emissions and Luxury Emissions, *Law and Policy* 15(1): 39–59.

Shue, H. (1994), After You: May Action by the Rich be Contingent upon Action by the Poor? *Indiana journal of Global Legal studies* 1:343-366.

Shue, H. (1995a,) Avoidable Necessity: Global Warming, International fairness and Alternative Energy. In *Theory and Practice,* NOMOS XXXV11, ed. Ian Shapiro and Judith Wagner Decew (New York: New York University Press), pp. 239-264.

Shue, H. (1995b). Equity in an International Agreement on Climate Change. In R. S. Odingo, A. L. Alusa, F. Mugo, J. K. Njihia, and A. Heidenreich, eds. *Equity and Social Considerations Related to Climate Change* (Nairobi: ICIPE Science Press): 385–92.

Shue H. (1996). Environmental Change and the Varieties of Justice. In Fen Osler Hampson and Judith Reppy, eds., *Earthly Goods: Environmental Change and Social Justice* (Ithaca, NY: Cornell University Press): 9–29.

Shue, H. (1999a,) Bequeathing Hazards: Security, Rights, and Property Rights of Future Humans. In M Dore, and Mount (Eds) *Global environment Economic: Equity and the Limits to Markets* (Oxford: Blackwell), pp. 38-53.

Shue, H. (1999b), Global Environment and International Inequality. *International Affairs* 75: 531–45.

Shue, H. (2014). *Climate Justice: Vulnerability and Protection.* Oxford: Oxford University Press.

Shue, H (2010), Deadly delays, saving opportunities: creating a more dangerous world In Gardiner S, Caney S, Jamieson D, and Shue H (Eds). *Climate Ethics: Essential Readings*, Oxford: Oxford University Press. pp.146-162.

Sinnott-Armstrong, W. (2010) It's Not *My* Fault: Global Warming and Individual Moral Obligation, in Gardiner S, Caney S, Jamieson D and Shue H (Eds). *Climate Ethics: Essential Readings*, Oxford: Oxford University Press. pp. 332-345

Slaughter, R. A. (1994). 'Why Should We Care for Future Generations Now.' *Futures*, 26(10): 1077–85.

Slovic, P. (2000). *The Perception of Risk*. London: Routledge.

Smith, P. T. (2018). Legitimacy and non-domination in solar radiation management research. *Ethics, Policy & Environment*, 21(3), 341–361.

Smith, A. ([1776] 1937). *An Inquiry into the Nature and Causes of the Wealth of Nations*, reprinted, New York: Random House.

Solar Radiation Management (2010): *The Governance of Research' by the Solar Radiation Management Governance Initiative* (SRMGI).

Spence A, Venables D, Pidgeon N, Poortinga W, Demski C.(2010) *Public Perceptions of Climate Change and Energy Futures in Britain: Summary Findings of a Survey Conducted in January-March 2010. Technical Report* (Understanding Risk Working Paper 10 - 01). Cardiff: School of Psychology. Available fromwww.understanding-risk.org

Stelzer, H. (2017).Justifying Climate Engineering? Dieter Sturma, Bert Heinrichs, and Ludger Honnefelder (Eds) *In Jahrbuch für Wissenschaft und Ethik*. Berlin: De Gruyter 21, 147-169.

Stemplowska, Z. (2008) What's Ideal about Ideal Theory, *Social Theory and Practice*, 34 (3): 319–40.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. New York, NY: Cambridge University Press.

Stern, N. (2015) *Why Are We Waiting: the Logic, Urgency, and Promise of tackling Climate Change.* Cambridge: Massachusetts Institute of Technology University Press.

Stevenson, H., and Dryzek, J. (2014). *Democratizing Global Climate Governance*. Cambridge, England: Cambridge University Press.

Svoboda, T. (2016). Aerosol geoengineering deployment and fairness. *Environmental Values.* 25 (1), 51–68.

Svoboda, T. (2017) *The Ethics of Climate engineering: Solar Radiation Management and Non-Ideal Justice.* New York: Routledge,

Svoboda, T. Irvine P, Callies D, and Sugiyama M. (2018) The potential for climate engineering with stratospheric sulfate aerosol injections to reduce climate injustice, *Journal of Global Ethics*, 14(3), 353-368.

Szerszynski, B., Kearnes, M., Macnaghten, P., Owen, R., and Stilgoe, J. (2013). Why Solar Radiation Management Geoengineering and Democracy Won't Mix. *Environment and Planning A: Economy and Space*, *45*(12), 2809–2816.

Tensen, E, Homann G. (2013) Implementing the precautionary principle for climate engineering in international law. *Carbon and Climate Law Review.* 2, 90–101.

Thaler, R. (1991), *Quasi rational economics*. New York: Russell Sage Foundation,

Thaler, R. (1994). Psychology and Savings Policies. *American Economic Review Papers and Proceeding*. 84 186–92.

Thaler, R and Sunstein, C (2008) *Nudge: Improving decisions about health, wealth and happiness* (New Haven: Yale University Press)

Thompson D. (1999). Democratic theory and global society. *Journal of Political Philosophy* 7(2):111–125.

Tilmes *et al* (2013) The hydrological impact of geoengineering in the Geoengineering Model Intercomparison Project (GeoMIP). *Journal of Geophysical Research*, 118:11036–11058.

Tollerfson, J. (2018). First sun-dimming experiment Will test a way to cool Earth. *Nature*, 563: 613-516.

Trisos C. H, Amatulli G, Gurevitch J, Robock A, Xia L , and Zambri B. (2018). Potentially dangerous consequences for biodiversity of solar geoengineering implementation and termination. *Nature Ecology and Evolution*, 2**,** 475–482

Turkaly C, Nicholson S, Livingston D. (2017) Establishing a Climate Engineering Research Clearinghouse, FCEA Papers and Reports. Washington, DC: Carnegie Endowment for International Peace.

Valentini, L. (2012). Ideal Vs. Non-ideal Theory: A Conceptual Map. *Philosophy Compass* 7 (9): 654–664.

Van Boven L, Ehret PJ, Sherman DK, (2018). Psychological barriers to bipartisan public support for climate policy. *Perspectives on Psychological Science* 13:492–507.

Van der Linden, S. (2015). The social-psychological determinants of climate change risk perceptions: towards a comprehensive model. *Journal of Environmental Psychology*. 41, 112–124

Vermeir, K. and Margocsy. D (2012) States of Secrecy. *British Journal for the History of Science*, 45 (2): 153-164.

Vogler, J. (1995). *The global commons: A regime analysis*. New York: John Wiley.

Wagner, G. and M. Weitzman. (2015). *Climate Shock: The Economic Consequences of a Hotter Planet*. Princeton University Press.

Waldron, J. (1999). *Law and Disagreement*. Oxford: Clarendon.

Warren, E. (1974). Governmental Secrecy: Corruption's Ally. *American Bar Association Journal* 60 (5): 550–552.

Weber, R.A. and Camerer C.F. (2006). "Behavioral experiments" in economics. *Experimental Economics*, 3 (9): 187–92.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, (39): 806–820.

Weisburd, S. (2015) Identifying moral hazard in car insurance contracts, *Review of Economics and Statistics* 97: 301–313.

Weiss, E B. (1992). In Fairness to Future Generations and Sustainable Development. *American University International Law Review*, 8(1): 19–26.

Whyte, K. P. (2012). Indigenous Peoples, Solar Radiation Management, and Consent. In C Preston (Eds). *Engineering the Climate: The Ethics of Solar Radiation Management*. Plymouth: Lexington Books, pp. 65-76.

Williams, B. (2005). *In the Beginning Was the Deed*. Princeton: Princeton University Press

Winter, G. (2011) Climate engineering and international law: last resort or the end of humanity? *Review of European Community and International Environmental. Law* 20, 277–289.

Wolff, J. (2010). Five types of risky situation, *Law, Innovation and Technology*, 2(2), 151–163.

Wolff, J. (2019) Fighting risk with risk: solar radiation management, regulatory drift, and minimal justice, *Critical Review of International Social and Political Philosophy*

Yitzhak, B, (2005), The Doctrine of Sufficiency: A Defence, *Utilitas* 17: 310–332.

Zürn, M. (2004). Global Governance under Legitimacy Pressure. *Government and Opposition* 39(2):260–287.

Zweifel P, and Manning W. (2000), Moral hazard and consumer incentives in health care, in A.J. Culyer and J.P. Newhouse (Eds.), *Handbook of Health Economics* (Elsevier, Amsterdam), Chapter 8.