Governing Prometheus
Ethical Reflections On Risk & Uncertainty In Solar Climate Engineering Research

Hofbauer, B.

DOI
10.4233/uuid:77b63c4a-85f7-4734-83f5-dd355b737191

Publication date
2024

Document Version
Final published version

Citation (APA)
Hofbauer, B. (2024). *Governing Prometheus: Ethical Reflections On Risk & Uncertainty In Solar Climate Engineering Research*. [Dissertation (TU Delft), Delft University of Technology].
https://doi.org/10.4233/uuid:77b63c4a-85f7-4734-83f5-dd355b737191

Governing Prometheus
Ethical Reflections On Risk & Uncertainty In Solar Climate Engineering Research

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of Rector Magnificus Prof. dr. ir. T. H. J. J. van der Hagen,
chair of the Board of Doctorates
to be defended publicly on Wednesday 19, June 2024 at 10.00 o’clock

by

Benjamin HOFBAUER
Master of Arts in Political, Economic, and Legal Philosophy, University of Graz, Austria
Born in Linz, Austria

This dissertation has been approved by the promotors.

Composition of doctoral committee:

Rector Magnificus	chairperson
Prof. dr. ir. B. Taebi	Delft University of Technology, promotor
Prof. dr. ir. I. R. van de Poel	Delft University of Technology, promotor
Dr. ir. Udo Pesch	Delft University of Technology, copromotor

Independent members:

Prof. dr. ir. H.W.J Russchenberg	Delft University of Technology
Prof. dr. ir. J.H. Kwakkel	Delft University of Technology
Prof. dr. S. Roeser	Delft University of Technology
Dr. P.-J Schweizer	Research Institute for Sustainability, Potsdam, Germany
Prof. dr. mr. ir. N. Doorn	Delft University of Technology, reserve member

Governing Prometheus

Ethical Reflections On Risk & Uncertainty In Solar Climate Engineering Research

Contents

Summary

This thesis explores the ethical challenges that a potential research program for solar climate engineering via Stratospheric Aerosol Injection (SAI) could incur. These ethical challenges are comprised of epistemic hurdles in relation to the research process, as well as societal questions of justice and the value of nature. The thesis proposes a variety of tools and approaches to assess and possibly govern the risks and uncertainties invoked by the research of SAI and its societal implications. The methodological approach is based mainly on ethical and philosophical analysis and reflection and the main findings take the form of discursive argumentation and normative reflection.

SAI is a form of climate engineering that seeks to reduce global warming by increasing the planet's reflection levels through the injection of reflective agents (aerosols) into the stratosphere. The mere potential of researching a technology that would actively intervene in the global climate is highly contentious and has led to passionate debates throughout the expert community. Designing a research process for such a polarizing technology such as SAI inevitably raises fundamental moral questions, wherein issues of global justice, democracy, the value of and humanity's relationship with nature, and the societal impacts of technological innovation all intersect. Given these far-reaching consequences, the thesis operates under the assumption that SAI is a highly disruptive idea and technology, that has the potential to challenge and undermine existing societal values and institutions. Accordingly, this work presents a range of philosophical modes of inquiry and assessments, in order to supply any proposed SAI research governance program with the necessary ethical considerations and frameworks.

The thesis is structured along four major inflection points, which form the individual chapters tied together through differing but interrelated research questions. What follows is an overview of the research questions, along with a brief description of how those questions were answered.

> Central Research Question: *In what ways should the desirability of researching Stratospheric Aerosol Injection be assessed, given its potentially highly disruptive effect on institutions, societal values and norms, as well as the physical environment?*

In order to answer this central research question, four subquestions are answered that constitute the four separate chapters of the dissertation.

> Subquestion 1: *What are the ethically pertinent uncertainties that arise in the context of an SAI research program, and how can they be accounted for?*

The first chapter introduces the central tenets of SAI research and its accompanying ethical issues. The main focus of the chapter is the introduction of a revised notion of "normative uncertainties", i.e., uncertainties that arise in multi-agential, value-pluralist and ambiguous decision-contexts. Making decisions surrounding the development and possible deployment of SAI within the context of climate change is, I argue, a prime example of such normative uncertainties. Participatory justice, recognition, as well as reflexive and adaptive governance approaches are presented as possible means to ethically account for these normative uncertainties.

> Subquestion 2: *What kind of risks and disruptions might a coordinated SAI research program incur on a systemic level, and how can they be ethically governed?*

In the second chapter, I connect SAI research governance to the concept of systemic risks. A central worry concerning SAI research is its potential to lead to reduced political, institutional, and societal support for radical mitigation. This risk is what some scholars have referred to as "mitigation deterrence" or "emissions abatement". In this sense, SAI research can be framed as a systemic risk to the energy transition and its underlying infrastructure. Systemic risks, as opposed to traditional risk concepts, are interdependent potential harms that occur in the context of modern societies' complex infrastructures and systems. They are trigger points that potentially lead to the breakdown of entire systems, with the housing market crash of 2008 and the Covid-19 Pandemic being two recent examples. The framing of SAI research's risks as systemic risks helps understand the central worry of mitigation deterrence better, while also allowing for the assessment and possible governance of such risks. In so doing, it also explores two possible supplements for the systemic risk framework in the form of value dynamism and meaningful participation and recognition. Through these ethical concepts developed throughout the philosophy and ethics of technology scholarship, the systemic risk framework becomes more holistic and able to deal with the justice and value questions related to SAI research, as well as New and Emerging Technologies (*NESTs*) in general.

Subquestion 3: *How can we assess the impact that an SAI research program might have on societal structures and institutions?*

On the basis of both the normative uncertainties and systemic risks of chapters one and two, the third chapter examines the ethical relevance of societal dynamics with respect to SAI research, specifically looking at the phenomenon of socio-technical lock-in. I tie the concept of lock-in to notions such as control, reversibility, and disruption through technological innovation, arguing that approaches such as the precautionary principle can give us a sense of what kind of innovation is ethically acceptable. However, any form of the precautionary principle comes with serious limitations when dealing with climate change, and can ultimately not answer fundamental questions as to what kind of societal values new and emerging technologies *should* enforce. On this basis, I conclude that the possibility of SAI research leading to some form of societal and institutional lock-in is in and of itself not morally problematic. The focus should be on what kinds of values are locked in, and to what extent these values are and should be reversible.

Subquestion 4: *How might an SAI research program affect complex societal values, such as sustainability?*

The fourth and final chapter of the thesis explores possible scenarios, or "vignettes", of techno-moral change put in motion by the implementation of an SAI research program. The chapter outlines different conceptions of the value of sustainability as a point of departure, exploring how SAI research might influence the conceptualization and societal relevance of that value. This approach seeks to bring together practices of techno-moral change and scenario-building, which does not aim at giving definitive predictions of value change and disruption. Rather, it tries to explore possible futures, giving scholars, as well as the public and policy-makers room to ask central normative questions about the kinds of societal changes we want to encourage, and how existing ways of seeing the world might be challenged through the development of e.g., SAI.

These four research questions highlight the need for dynamic and embedded ethical assessment frameworks for the research governance of SAI and beyond. The central argument follows the idea that technology assessment needs to be synchronous to the research and development phase of any new and emerging technology, and that the interaction between the innovation process and societal values is a foundational aspect of ethical research governance. Further, understanding and accounting for the moral context of the risks and uncertainties involved in such processes is paramount, in order to both act with precaution and humility, but also see the potential opportunities for positive changes and disruptions of entrenched value systems through innovation. Beyond the scope of SAI research governance, this thesis is also of generic relevance for the ethical and philosophical assessment of emerging technologies and innovation, the intersection between environmental ethics and the ethics of technology, as well as risk ethics and the dynamics of socio-technical systems.

Glossary

BECCS	Bio Energy Carbon Capture & Sequestration
CDR	Carbon Dioxide Reduction
CE	Climate Engineering
CO2	Carbon Dioxide
DACCS	Direct Air Carbon Capture & Sequestration
GHG	Greenhouse Gas
IPCC	Intergovernmental Panel on Climate Change
NASEM	National Academies of Science and Medicine
NESTs	New and Emerging Sciences and Technologies
NET	Negative Emissions Technology
NU	Normative Uncertainty
SAI	Stratospheric Aerosol Injection
SRM	Stratospheric Aerosol Injection
TMC	Techno-Moral Change

Prologue: The Power of Technology and Science

And our singing shall build
In the void's loose field
A world for the Spirit of Wisdom to wield;
We will take our plan
From the new world of man,
And our work shall be called the Promethean.

~ Prometheus Unbound, Act IV,
Percy Bysshe Shelley

Science, and its application through technology and engineering is an incredibly powerful means to shape not just our environment, but the way in which we perceive the world. There is a deeper truth in the saying “If you have a hammer, everything becomes a nail.” By opening up new possibilities of how to engage with the world, the hammer also changes our surroundings: A previously uninteresting nail becomes a useful element for construction, a previously impenetrable wall becomes a potentially breakable object. The fact that science, technology, and engineering are not just means to ends, but actively co-create our reality is a central insight from the philosophy of technology, as philosophers such as Martin Heidegger, Bruno Latour, Gilles Deleuze, or Hans Jonas have reflected on. Importantly, the kinds of questions we ask through science; the kinds of technologies humanity chooses to develop are also manifestations of the values, norms, and beliefs that humanity holds about how the world *should* be (Midgley 2006). This raises two ethical questions, one descriptive, one normative. Descriptively, we can ask what kind of values, norms, and beliefs existing and developing technologies represent, and what the underlying reasons for their development are. Normatively we need to consider whether or not a certain line of inquiry - scientific or technological - should be pursued in the first place. Is the action, value, or purpose good, which that technology would facilitate? Does it infringe upon justice, equity, well-being?

We find references to these fundamental questions of ethics and technology throughout (Western) history. Often, literature and mythology present us with warnings against the hubristic tendency of humanity to strive beyond its capacities. When Icarus sought to achieve the heavenly power of flight, or Dr. Frankenstein obsessed over the godly power to create life, they were both met with the crude reality of human limitation. They overestimated their control, skill and put confidence into a technology they did not properly grasp, and they were humbled as a consequence. However, mythology in particular also offers us at least one example of the emancipatory force of technology, and of how science can serve as a means to break the chains of the limited human condition. Such is the story of Prometheus, the titan who brought fire, i.e., knowledge manifested, to humanity in spite of the Godfather Zeus' commands. While there are of course numerous possible interpretations of this age-old tale, following the title and introductory lines of the poet Percy Bysshe Shelley, I want to focus on a specific perspective on the Prometheus Saga. Namely that science and technology are the key to human development, to breaking the shackles of natural boundaries. I believe this core-idea that the romantics, German idealists and the "Sturm und Drang"-Poets throughout Germany and England have relied on, can be fruitfully tied to many of the issues surrounding the burning questions about technological solutionism, techno-optimism and pessimism, deep ecology, eco-modernism, etc.

Romanticism was both an embrace of the development of human capacities through science, as well as a warning and struggle against the increasing hyper-rationalism of the 19th century (Ferber 2010). Similarly, I see my work as trying to find ethical ways of evaluating and developing technologies that can increase overall well-being, rather than merely benefiting a wealthy few and disregarding its societal and environmental consequences. Solar climate engineering, as a technological intervention to potentially ease some of the anthropogenic climate harms we are confronted with, is a prime example of a technology that wanders the tight rope between reasonable necessity and irredeemable hubris. This thesis is an exploration of how we can govern the knowledge production related to solar climate engineering, and what it takes for the Promethean enterprise of researching Stratospheric Aerosol Injection to be done as an ethical process, guided by principles of justice and being proactive in its account of risk and uncertainty through systems-based thinking.

I. Introduction: Climate Engineering & the Climate Catastrophe: Some Context

Radical mitigation and continuous adaption strategies are vital in humanity's struggle with climate change. Handling climate change effectively, sustainably, and equitably can only be done with the mitigation and adaptation as core policy tenets. At the same time, given the scale of the challenge, these two strategies alone may not suffice to stave of the most catastrophic climate impacts within the 21st century. As of yet the radical mitigation necessary to overcome climate change is woefully insufficient and adaption can only be part of the solution. This is why the idea of directly influencing the climate through technological means and engineering has gained some prominence in the past two decades. So called "geoengineering" or "climate engineering" (CE) may be considered to play a central role in humanity's overcoming of anthropogenic climate change.

Since the 2009 report of the Royal Society one can see the forms of CE being split into two separate categories: Solar Radiation Management or Modification, and Carbon Dioxide Removal (Shepherd et al. 2009). Carbon Dioxide Removal (CDR) describes any active intervention that leads to the reduction of atmospheric CO_2. Consequently, CDR ranges from technological tools such as Direct Air Carbon Capture & Sequestration (DACCS) and Bio Energy Carbon Capture & Sequestration (BECCS), to nature-based approaches such as ocean fertilization and reforestation.

While CDR approaches invoke important ethical issues in their own right, for the scope of this work I focus on Stratospheric Aerosol Injection (SAI), which is a form of SRM. The implementation of SAI would entail continuously deploying aerosols into the stratosphere in order to increase the planets reflection level or albedo. Based on climate modelling and empirical observations after a volcanic eruption, the assumption is that increasing the earth's reflection levels would lead to a cooling of the global surface (Caldeira, Bala, and Cao 2013). According to various specific models, the deployment of SAI is estimated to mitigate some of the more severe climate risks. It should be noted that there is a near unequivocal consensus within the modelling community of SAI that it does not represent a silver bullet solution towards anthropogenic climate change. Rather, it is considered to be part of a portfolio solution towards combatting climate change, in concert with radical mitigation efforts, adaption, and methods to remove carbon dioxide from the atmosphere. An illustrative example of this is the idea of "shaving the peak” off of the most severe warming of the climate (Tilmes et al. 2020). In times of peak emissions, SAI could be used to mitigate some of the graver risks by cooling the planet despite overshooting emissions. After emissions have been acceptably reduced, the deployment of SAI is phased out. This is just one example of how SAI is imagined to be used.

The call to at least investigate a seemingly hubristic technology such as SAI is justified through the dire urgency climate change invokes. The argument from hubris mainly stems from the idea that humanity should not meddle with things it vitally depends on without due caution. Since we barely understand all the underlying dependencies and systematic interactions of the climate system and climate change, adding another risk-factor in the form of SAI to the mix is naïve at best, and an injustice at worst so the claim goes (Hamilton 2013). The urgency to investigate SAI nonetheless comes from two central interdependent dangers that the current climate trajectory entails. One, crossing certain climate thresholds could lead to irreversible shifts in various global ecosystems and weather patterns. Examples of such shifts are the switching of rainforests into savannahs, the melting of permafrost, changes to the Atlantic Meridional Overturning Circulation (AMOC), or the complete loss of arctic sea ice in the northern summers (Good et al. 2018; Lenton et al. 2008; Rockström et al. 2009).

The second danger is ever-increasing weather extremes in the form of wildfires, droughts, floods, storms, etc. that climate change produces. While humanity is already in the middle of the climate catastrophe unfolding, the crossing of said thresholds would further cause irreversible changes that would increase climate dangers and lead to an increasingly unpredictable future for humanity on this planet. For this reason, immediate cessation of carbon dioxide emissions alongside robust adaptation measures are paramount to set favourable conditions for the continuous existence of human societies, and to reduce the rising rates of extinction and degradation of existing flora and fauna. Nonetheless, mitigation and adaptation might need to be complemented through climate engineering approaches such as CDR, SRM and potentially SAI.

Urgency and calls beyond mitigation and adaptation are further justified on the basis of the climate system’s inertia. The climate system lags behind in its reaction to the increased levels of atmospheric carbon dioxide. Consequently, merely abating emissions is not enough to meet the 1,5/2 degrees Celsius targets set by the Paris Climate Accord. As the Intergovernmental Panel on Climate Change states (IPCC 2021), keeping global temperatures beneath 2 degrees Celsius by

2100 and after, will require some form of reducing atmospheric CO_2, relying on natural carbon sinks (forests, the ocean, etc.) as well as technological solutions (e.g., Bio Energy Carbon Capture & Sequestration and Direct Air Carbon Capture &Sequestration). A climate future below 2 degrees thus already relies, in part, on technological solutions through Carbon Dioxide Reduction technologies and concepts (Brack and King 2021).[1] It is unclear whether sensible CDR technologies that can meet the demands of sheer scale that a globally effective reduction of atmospheric carbon would require are currently feasible for deployment. With the potential crossing of various climate tipping points on the horizon however, humanity is in a position of relentless urgency to act decisively and quickly, two attributes ill-suited to describe the relatively stable, yet cumbersome socio-political systems, institutions, and structures of the 20th century (Dryzek and Pickering 2019).

It is within this framing of urgency, pressure, and impeding catastrophe that some researchers are suggesting additional means of climate action through the active modification of incoming solar radiation, or SRM for short. Similar to CDR, SRM is a form of climate engineering, i.e. an intentional manipulation of the climate through technological means (Keith 2013; Shepherd et al. 2009). Like CDR, SRM should also be understood as a principle of intervention rather than a technology in and of itself. Either principle focuses on one of the central levers of the global climate, namely the storage of energy through greenhouse gases (CDR) on the one hand, and the earth's main energy source, solar radiation (SRM) on the other hand. An important difference between SRM and CDR is that SRM is purely symptomatic in its treatment of climate change, and is therefore less of a viable solution, and more of a temporary relief. It is particularly the second lever in the form of SRM that is the source of polarization throughout the intersection of governance, ethics, and environmental sciences.

I.II Stratospheric Aerosol Injection: A History of Ideas & The Current Discourse

From a climatological perspective, SAI "merely" alleviates one of the many threats of climate change, namely global warming. Other issues such as the enormous loss in biodiversity, ocean acidification, environmental degradation, and generally unsustainable consumption of resources remains unaddressed by SAI. In short, climate engineering with SAI would be a symptomatic

[1] The models implying the use of NETs in the future rely on continuous economic growth, which is a contentious assumption in and of itself. However, given the climate system's inertia there is no reason to believe that immediate cessation of CO2 emissions would be enough to fend of irreversible harms to the planet.

treatment for the chronic ailment of a carbon-dependent society. The ethical debate on its implementation ranges from dangerous and irresponsible (Gardiner 2009; Hulme 2014), to dangerous but necessary (Horton and Keith 2016; Keith and Irvine 2016).

Following Nobel laureate Paul. J. Crutzen's editorial essay on a potential urgency to deploy SAI, a lively debate and tentative, model-based research surrounding the topic has started (Crutzen 2006). SAI is being analyzed in terms of its potential costs (Moriyama et al. 2017; Smith and Wagner 2018), its modelled impacts on the climate (Kravitz et al. 2011) and potential implementation. In terms of ethical implications and desirability, the potential of thinking about, researching, deploying, and governing SAI has fostered a somewhat lively community with stances for or against (for a broad overview, see Preston 2012). By following and extrapolating from the modelled outcomes from potential SAI deployment, the question of whether or not it may be morally permissible to "engineer the climate" is contentiously discussed (Biermann et al 2022; Wieners et al 2023).

The debate surrounding the admissibility of SAI and its research has been developing throughout the past decade and is slowly gaining traction in the policy arena as well. As an example, the international group of policymakers, political leaders, and researchers convening as the "Global Commission on Governing Risks from Climate Overshoot", or "Climate Overshoot Commission" for short, has recently published a report that highlights the need to also actively research Solar Radiation Modification approaches.[2] While the group urges against any form of deployment of SRM, its explicit support for SAI research is significant, given the way in which SAI has usually been left out of serious policy discussions so far. The Commission, consisting of members from the global north and south, focuses on the likely scenario of humanity not reaching the Paris 2015 goal of staying below an increase in global mean temperature of 1,5 degrees Celsius. Accordingly, they recommend increased mitigation, especially from industrialized nations, adaptation for both immediate climate disasters, long-term climate change impacts, and indirect effects of climate change such as migration shifts. The group also recommends more investment and research on Carbon Dioxide Removal, and most controversially more research into Solar Radiation Modification.

[2] https://www.overshootcommission.org [last accessed November 2023]

While these recent developments seem unprecedented, the idea to actively change the climate or weather through some form of human intervention is not completely novel. A number of scholars have traced this desire to shape the climate back to at least the cold war era (Stilgoe 2015; Morton 2015; Baskin 2021). Jeremy Baskin for example recounts the debate surrounding climate engineering in various phases throughout the 20th to 21st century (Baskin 2021, 28ff.), and highlights how the Cold War arms race entrenched a militarized approach towards science and NESTs in general, and climate engineering in particular as a form of "mastery" over nature. While this view slowly dissipated, ultimately dropping climate engineering for other endeavours and shifts in both the economic situations of many countries involved, as well changes in the political positions of the citizens, its influence is undoubtedly nowadays felt during climate engineering's revival.

Even so, it may be argued that the very concept of exerting power and control over nature is the fundamental premise for (Western)[3] conceptions of societal progress. Facing nature as something to be handled and controlled can for instance be traced back to ancient Greek philosophers and mathematicians such as Archimedes. The quote attributed to him "give me the place to stand I shall move the earth", can be read as a resounding conviction about the ultimately limitless power of science and technology (Stilgoe 2015). In a similar vein, the philosophy of science scholar Mary Midgley discussed the crucially mythical aspect of science as a post-religious means of salvation for modern societies, wherein "neo-MAN" becomes god, shaping themselves and their surroundings according to their will. (Midgley 2001, 26). Science, and its application in the form of technology become the emancipatory power for humanity to create their own fate. From this perspective, an interest and push for climate engineering is not just explainable, but rather it becomes the inevitable conclusion to humanity's struggle with the environment.

To what extent ever-increasing, seemingly endless forms of innovation are an inevitable feature of (Western) forms of doing science and understanding the world, is an important question that merits attention when considering such technologies. Its importance is particularly central to any possible research process of NESTs such as SAI. Research and innovation processes are not socio-politically neutral, but rather value-laden undertakings. They are representative or run

[3] Please note that I add "Western" in parenthesis here to indicate that my claims are to be taken in the context of a Western upbringing and modes of thinking. While I am in not qualified to assess this, it is possible that these claims apply to a non-Western context as well.

counter to existing world views, and thus possibly reinforce or challenge existing societal structures. With regard to SAI research then, the question is how a research governance process can be developed that accounts for the kinds of justice concerns that inevitably form part of a technology that is envisioned to be deployed on a global scale, affecting the entire planet over generations to come.

I.III SAI Research and the Ethics of Socially Disruptive Technologies

SAI undoubtedly raises numerous ethical questions, given its wide-ranging implications on both a physical and even more on a societal level. This raises a generic question with regards to new and emerging technologies in general, namely whether certain technologies should be researched or developed, whether certain investigations should be undertaken in the first place and how much resources should go into them. These are all questions that fundamentally require moral arguments. While the ethical challenges have particular weight in the case of SAI, an invariably global technology, ethical reasoning about technology spans far beyond climate engineering.

The Ethics of Socially Disruptive Technologies[4] (*ESDiT*) research program that this thesis on the ethics of governing solar climate engineering research forms part of thus seeks to deal with any form of "socially disruptive technologies". Socially disruptive technologies produce both the fear of impending doom through technologies humanity does not fully grasp, and the hope of achieving future technological utopias. In line with this ambiguity, the ESDiT research program investigates an array of new and emerging technologies, such as nanotechnology, Artificial Intelligence, biotechnology, social media, and climate engineering among others. All these technologies share the potential of being socially disruptive, as well as disrupt ethical concepts and theories (van de Poel et al. 2023). They challenge traditional concepts of privacy (big data and social media), what it means to be a human being (robotics and AI; biotechnology) gender roles and norms (artificial gestation), humanity's relationship with nature (climate engineering and artificial meat), and many other previously considered stable ways of understanding the world.

As the polarizing debate surrounding SAI research hints at, the very idea of solar climate engineering is disruptive in numerous ways. On a societal level, SAI research invokes

[4] https://www.esdit.nl/ [accessed October 2023]

fundamental questions of global, environmental and intergenerational justice, as well as unprecedented hurdles for fair governance and policy. Relatedly, the ethical and philosophical implications of seeking to in some way engineer the climate raise profound questions about the way humanity relates to itself, to nature and to the non-human world.

I. IV State-of-the-Art SAI Research Governance and Challenges

With both the urgency of climate change and the potential disruptiveness of SAI in mind, it is understandable that there has been a growing amount of scholarship on how to design the research and governance process of Solar Climate Engineering, and SAI in particular. Two recent official reports within the policy-making arena have been influenced by and influential for these efforts, namely the 2021 report "Reflecting Sunlight. Recommendations for Solar Geoengineering Research and Research Governance" by the National Academies of Science's (NASEM 2021), and the "One Atmosphere: An independent expert review on Solar Radiation Modification research and deployment" report from the United Nations Environment Programme from 2023 (UNEP 2023).

A central observation that these two reports share is the ubiquity of risk and uncertainty embedded in research process of SAI. At a glance, this is an unavoidable side-effect of dealing with climate change, as the climate and the weather are themselves complex systems that rarely allow for precise, stable, and long-term predictions. How climate change will play out exactly, and what its precise impacts will be is often only assessable after the fact, i.e., after a catastrophic flood, drought, or storm has occurred, or an ecosystem has already entered a phase of irreversible collapse or change. However, in the case of climate engineering through SAI, the complexity increases even further due to the addition of an active sociotechnical intervention into the climate system. An SAI research process cannot solely rely on natural science predictions and models about global and regional weather impacts. It must also take into account the sociopolitical and ethical implications. As much as humanity shapes such a research program, SAI research will also shape humanity.

The authors of the mentioned reports are keenly aware of this complexity and the need to account for the sociotechnical risks and uncertainties involved in researching SAI. Thus, central governance tenets such as public engagement and communication (NASEM 2021, 171), the need for inter- and transdisciplinary research (UNEP 2023, 25; NASEM 2021, 142), as well as clear cut “exit-ramps” (NASEM 2021, 9) that allow for the suspension and cancellation of the research program all form part of these reports.

Undoubtedly, the reflective and thorough nature of these reports is also owed to the active scholarship on research governance and responsible innovation, and the ethics of technology. For example, the Science, Technology, and Society scholars Stilgoe, Owen, and Macnaghten describe four central tenets of responsible innovation, namely anticipation, reflexivity, inclusion, and responsiveness (2013, 1570). Crucially, responsible innovation and its accompanying fields such as anticipatory governance, technology assessment, and value-sensitive design, all seek to explicitly introduce moral considerations into the scientific process, rather than having ethics become an afterthought to innovation. Further, the ethical governance of innovating responsibly, is driven by a forward-looking understanding of responsibility (van de Poel 2011). This entails that researchers and those overseeing research programs ensure “that responsible choices can be made in the future, through anticipating and gaining knowledge of possible consequences and building capacity to respond to them.” (Stilgoe et al. 2013, 1579)

Responsible innovation through ethically reflected governance processes opens up space for deliberation on what kinds of values, norms, or moral considerations should form part of such a research program. Thus, diverse conceptions of justice, in the form of distribution, procedure, and recognition are regarded as pivotal for the responsible development of SAI (for an overview, see Preston 2016). Scholarship shows an evolution from the hypothetical justifiability of SAI deployment to concrete recommendations for a fully-fledged SAI research program. A key reference point here is the 2009 piece by Alan Robock and colleagues that seeks to discuss the possible up and downsides to SAI deployment (Robock aet al. 2009). Further important research endeavours with regards to SAI governance and its risks was the report on climate engineering published by the Royal Society in 2009 (Shepherd et al. 2009), as well as the recommendations coming on the back of the 2010 Asimolar Conference on climate engineering (Asimolar Scientific Organizing Committee, 2010). This scholarship prompted all kinds of responses, particularly from ethicists and philosophers who were often highly critical of SAI research (see for example, Gardiner 2009; Szerszynski et al. 2013; Fragnière and Gardiner 2016). Worried that SAI would reduce the political incentive for radical mitigation, as well as not tackle the underlying societal issues related to climate change, the critical voices surrounding SAI culminated in a 2022 call to put a moratorium on SAI deployment (Biermann et al. 2022).

At the same time, the drastic developments surrounding climate change and its ever-more catastrophic impact has moved some scholars towards a tacit acceptance of at least developing research programs. Gardiner and Fragnière for example have developed their own set of ethical principles, building on Robock et al. 2009 and the Asimolar Principles (Gadiner and Fragnière 2018). The "Tollgate Principles" (Gadiner and Fragnière 2018) seek to expand on previous frameworks by building a more substantive ethical understanding of the need to govern SAI research on the basis of justice and legitimacy concerns. Further, some scholars have come out in favour of establishing clear guidelines and principles on how SAI research could be governed, seeking to expedite SAI research processes while also bringing best practice and moral considerations to the table (Wieners et al. 2023). It is from this background of polarization and reluctant acceptance that my investigation departs.

I.V Objectives & Research Questions

This dissertation is premised on two, co-dependent points of inquiry that together form the central research question:

> RQ: *In what ways should the desirability of researching Stratospheric Aerosol Injection be assessed, given its potentially highly disruptive effect on institutions, societal values and norms, as well as the physical environment?*

This research question runs through the entire thesis and is examined throughout numerous perspectives, namely techno-moral imagination, risk ethical considerations, participatory and adaptive governance approaches, as well as concepts found in justice theory and the ethics of technology. Importantly, exploring this RQ does not result in a substantive answer that declares SAI research to be "good" or "bad". Instead, the thesis presents various procedural solutions to the various ethical issues arising from the design and potential implementation of a given SAI research governance proposal.

In short, SAI research invokes complex, interdependent risks and uncertainties that arise in the context of moral ambiguity and ambiguous data. For the moral assessment of these risks and uncertainties, a meaningful participation processes, ethical reflection, and techno-moral scenario building become central. Importantly, the thesis builds in many ways on existing SAI research governance proposals and seeks to deepen the concerns surrounding a research program's unintended outcomes (Stilgoe et al. 2013; Bellamy 2016; NASEM 2021; UNEP 2023).

Answering the main Research Question in depth invites a set of 4 subquestions, answered in the separate chapters of the thesis.

> Subquestion 1: *What are the ethically pertinent uncertainties that arise in the context of an SAI research program, and how can they be accounted for?*
>
> Subquestion 2: *What kind of risks and disruptions might a coordinated SAI research program incur on a systemic level, and how can they be ethically governed?*
>
> Subquestion 3: *How can we assess the impact that an SAI research program might have on societal structures and institutions?*
>
> Subquestion 4: *How could an SAI research program affect complex societal values, such as sustainability?*

Answering these subquestions gives a nuanced perspective on the ethical challenges raised by SAI research governance, while also providing possible avenues to deal with these challenges.

The first subquestion explores what I refer to as "normative uncertainty". Normative uncertainty arises in the context of multi-lateral decision-situations, wherein a plurality of values need to be reconciled under factually and ethically ambiguous circumstances. Making decisions about the research and possible deployment of SAI gives rise to such normative uncertainty, with the technology having inherently global impacts, its use spanning over at least two generations of human beings, and its exact outcomes being reasonably debatable until fully deployed. The ethical governance of SAI research needs to come with robust participation procedures that take into account the potential historical inequities among participants, as well as adaptive governance procedures that allow for a dynamic and iterative decision-making process.

Understanding the ethical relevance of uncertainty arising in the context of SAI research opens up space to delve further into the uncertain impacts related to such a research program. In particular, the second subquestion deals with the risk of mitigation deterrence, one such potential unexpected outcome. Namely, mitigation deterrence describes the risk of SAI research leading to reduced efforts in mitigation, as politicians and parts of the public might use the technology as a potential excuse to invest less into radical emissions reduction. The second chapter frames mitigation deterrence as a systemic risk. Systemic risks, as opposed to conventional risks, arise in the context of complex, interdependent systems and structures. Their effects are non-linear and often manifest themselves in the form of tipping-points. As a conceptual tool, systemic risks provide an insightful framework for the exploration of SAI research risks given its systemic approach that accounts for the complex interactions of different systems and structures, such as the energy infrastructure, policy frameworks, and institutional settings. At the same time, the framework requires some philosophical supplements in order to better identify the ethical implications of a NEST such as SAI. Consequently, the second chapter offers insights from the philosophy and ethics of technology to supplement the systemic risk framework with an ethically reflected conception of participatory justice, as well as a dynamic account of the value change and technological development.

With these risks and uncertainties in mind, the third subquestion focuses on the potentially disruptive societal impact of SAI research, and how that impact can be ethically assessed. The three focal points for this inquiry are disruption, societal control, and lock-in. Most importantly, SAI research can cause a form of societal disruption, wherein existing norms and values change, or new sociotechnical imaginaries arise as a consequence of the technology's development. This raises the question of ethical desirability, i.e., whether this disruption can be morally justified. The central argument of the third chapter is that the admissibility of this disruption is highly dependent on how reversible its impacts are, in line with approaches such as the precautionary principle. At the same time, reversibility alone is not a sufficient criterion for moral justification, but necessitates clarity in terms of what ought to be reversible (values, institutions, infrastructure), and to what degree.

The fourth and final subquestion serves as a set-up for a form of techno-moral scenario building. In this sense, trying to understand the potential impacts that SAI research may have on the value of sustainability is not an exercise in prediction or modelling. Instead, chapter four provides a space for deliberation and debate. In so doing, it presents two possible vignettes of how SAI can either be seen as a moral catalyst for sustainability efforts or undermine the very concept to the degree of irrelevance. In short, SAI could lead to a concerted global effort in radical emissions reduction since it underscores the importance of cooperation and urgency: Humanity has come to a place where engineering the climate became necessary, and this served as a global wake up call. At the same time, SAI could lead to the further entrenchment of existing laissez-faire attitudes toward climate change and ecological catastrophe: Why be sustainable in a world where we can engineer the climate? Neither of these vignettes are intended to be predictions of the future, but rather open up space for moral deliberation and imagination, which in turn could guide participatory procedures and policy-making decisions.

I.VI Overview of chapters

The dissertation's main body consists of four chapters, which each answer one of the subquestions respectively. Three of the chapters are published in peer-reviewed journals, and one is currently under review. Their titles and abstracts are presented below. Chapter 1 and chapter 4 have been published and their references are provided as footnotes (Hofbauer 2023a, Hofbauer 2022), and chapter 2 has been accepted for publication (Hofbauer 2023b). Chapter 3 is currently under review.

Chapter 1: *Normative Uncertainty in Solar Climate Engineering Research Governance*[5]

This paper explores what kind of uncertainty a research program governing solar climate engineering through Stratospheric Aerosol Injection (SAI) needs to account for. Specifically, it tries to answer two central issues with regards to SAI research and its ethical evaluation: One, what irreducible uncertainties remain throughout the decision-process, and two, how do these remaining uncertainties affect the ethical evaluation of SAI research? The main assumption is that decisions on SAI research governance will be made under normative uncertainty, i.e., situations under irreducible knowledge-constraints that arise in concrete, practical decision-contexts. These decision-contexts are multi-lateral and empirically ambiguous, and the decision-makers need to reconcile a plurality of values. While normative uncertainty complicates the ethical evaluation of policy decisions, I argue that moral considerations can be accommodated through the inclusion of recognitional, participatory justice approaches, as well as adaptive and anticipatory governance methods.

[5] Hofbauer, Benjamin. "Normative Uncertainty in Solar Climate Engineering Research Governance." *Ethics, Policy & Environment* (May 26, 2023): 1–20. https://doi.org/10.1080/21550085.2023.2216148.

Chapter 2: *Putting Cats Back Into Bags. Disruption, Reversibility, & Lock-in in Solar Climate Engineering*[6]

This paper critically examines the normative guidance as well as practical feasibility of reversibility as an ethical criterion for New and Emerging Technologies, with a focus on solar climate engineering through Stratospheric Aerosol Injection. Specifically, concepts such as lock-in, control, and disruption, all relate to some degree to the idea that reversing unwanted technological impacts is both morally desirable and prudent given the risks and uncertainties NESTs introduce. At the same time, the exact implications of what should be reversible remains unclear, whether it be societal institutions, value systems or collective worldviews, physical and non-physical infrastructure, etc. The claim is that while a useful tool for the investigation of new technologies, reversibility needs to be strictly qualified and contextualized, if it is to serve as a criterion for ethical and practical guidance.

Chapter 3: *Systemic Risks and Solar Climate Engineering Research. Integrating Technology Ethics Into The Governance of Systemic Risks*[7]

The paper explores how the framework of systemic risks can help govern the risks imposed through solar climate engineering research. The central argument is that a systemic perspective of risk is a useful tool for analysing and assessing the risks imposed through Stratospheric Aerosol Injection (*SAI*) research. SAI is a form of climate engineering that could cool the planet by enhancing its albedo through the injection of aerosols into the stratosphere. Researching such a technology creates is systemic risks with a strong sociotechnical component. This component consists of the potential societal harm that a developing or new technology might cause to existing norms, values, institutions, and politics. The systemic risk framework is a valuable heuristic for this case, given the complex interdependencies of societal systems, infrastructures, markets, etc. At the same time, systemic risk scholarship has so far not dealt with the ethical issues arising in its own framework, especially with regards to the ethical aspects of technological risks. Consequently, this article seeks to supplement the systemic risk governance

[6] Under review [date of submission: 29.08.2023]

[7] Accepted in *Journal of Risk Research*, *https://doi.org/10.1080/13669877.2023.2288010*.

framework with insights from technology ethics. Specifically, the paper offers an ethically reflective conception of societal value dynamism and stakeholder engagement and participation, tying it to existing systemic risk governance approaches. A potential research program for Stratospheric Aerosol Injection serves as a case study.

Chapter 4: *Techno-moral change through solar geoengineering: How geoengineering challenges sustainability*[8]

This article brings a new perspective to the ethical debate on geoengineering through stratospheric aerosol injection (SAI), incorporating the emerging techno-moral change scholarship into the discussions surrounding sustainability. The techno-moral change approach can help us understand different ways in which technology might shape society. First, it helps highlight how values and norms are interrelated. Second, it shows that techno-moral change can happen even if the technology is in no way realized. Through the introduction of two techno-moral vignettes, two diametrically opposed ways in which SAI forces us to rethink sustainability and our relationship with nature are suggested. SAI could lead to a situation of entrenchment, wherein sustainability as a norm is undermined, or transformation where the necessity of acting according to sustainability is highlighted.

[8] Hofbauer, Benjamin. "Techno-Moral Change through Solar Geoengineering: How Geoengineering Challenges Sustainability." *Prometheus*, June 1, 2022. https://doi.org/10.13169/prometheus.38.1.0082.

II. Chapter 1: Normative Uncertainty in Solar Climate Engineering[9]

This paper explores what kind of uncertainty a research program governing solar climate engineering through Stratospheric Aerosol Injection (SAI) needs to account for. Specifically, it tries to answer two central issues with regards to SAI research and it's ethical evaluation: One, what irreducible uncertainties remain throughout the decision-process, and, two, how do these remaining uncertainties affect the ethical evaluation of SAI research. The main assumption is that decisions on SAI research governance will be made under normative uncertainty, i.e. situations under irreducible knowledge-constraints that arise in concrete, practical decision-contexts. These decision-contexts are multi-lateral and empirically ambiguous, and the decision-makers need to reconcile a plurality of values. While normative uncertainty complicates the ethical evaluation of policy decisions, I argue that moral considerations can be accommodated through the inclusion of recognitional, participatory justice approaches, as well as adaptive and anticipatory governance methods.

II.1.1 Introduction

Given the dire outlook and urgency of climate change, technological solutions are being discussed by researchers and policy makers alongside proposals for adaption and radical mitigation. A particularly impactful, potential climate technology is the concept of climate engineering, i.e. intentionally manipulating the climate system through technological means (Keith, 2001; Shepherd et al., 2009)

One way of conducting climate engineering could be through so called Stratospheric Aerosol Injection (henceforth *SAI*). SAI would artificially slow down the warming of the planet by increasing the stratosphere's reflection levels. Increasing the planet's albedo through SAI entails spraying aerosol particles into the stratosphere, which in turn reduces some of the incoming sunlight (Crutzen, 2006). However, the ethics surrounding SAI are contentious and polarizing

[9] This chapter is published as: Hofbauer, Benjamin. "Normative Uncertainty in Solar Climate Engineering Research Governance." *Ethics, Policy & Environment* (May 26, 2023): 1–20. https://doi.org/10.1080/21550085.2023.2216148.

raising issues of distributive, participatory and historical (in) justice, the relationship between human beings and nature, risk-ethical issues, etc. (Pamplany et al., 2020)

Although SAI receives marginal but continuous attention within the ethics literature (for an overview, e.g. Preston, 2013), a focus on the uncertainties in the debate is still lacking. Specifically the implications and limitations uncertainty causes for the ethical assessment of SAI are in need of investigation.

While the very intention of SAI is to have a global effect, i.e. reducing the rate of warming, many of its effects, such as regional impacts, as well as its societal consequences are shrouded in considerable, at times irreducible uncertainty. Irreducible in this case entails that the uncertainty cannot be resolved through more empirical research or ethical analysis. Rather, the decision-process needs to take place under uncertainty. Such irreducible uncertainty in SAI research poses a challenge for its ethical assessment, which opens up two central questions. Namely, what kinds of irreducible uncertainties remain, and, how can these remaining uncertainties be ethically accounted for in the decision- making process?

In order to answer these two questions, I propose an adapted notion of normative uncertainty, that builds on previous notions of uncertainty as discussed within the scholarship surrounding precaution, decision theory and metaethics, as well as Responsible Research and Innovation. I argue that normative uncertainties are epistemic states wherein decision-makers have no unequivocal answer to a given question due to both the empirical circumstances of the situation as well as its ethical context. They arise in multi-lateral decision-situations that are empirically ambiguous, and wherein the decision-makers need to reconcile a plurality of different values. Normative uncertainties are not mere value disagreements, but rather challenge the adequacy and applicability of any given value due to the empirically uncertain circumstances (e.g. a lack of predictability and understanding of the empirical facts surrounding a given decision).

Importantly, the notion of uncertainty presented here is contextualized within the complexity of climate policy-making. This entails embedding decision-making under uncertainty within a multi-agential decision-space, representing numerous and diverse stakeholders, which in turn adhere to diverse ethical convictions. Decisions surrounding SAI, an inherently global intervention, can only be made sense of within this context of complex, multi-agential and global consideration.

Section I contextualizes the piece within the state-of-the-art SAI research governance literature, specifically focusing on the National Academies' 2021 report on solar climate engineering research governance. On this basis, Section II lays out how SAI research relates to the concept of precaution and uncertainty, arguing that precaution merely focuses on the empirical aspect of uncertainty. This focus on empirical uncertainty ignores the way in which moral disagreement and value pluralism lead to moral uncertainty. Consequently, the third section discusses ways of implementing moral uncertainty, found in decision theory and metaethics, into the broader, concept of normative uncertainty, accounting for both moral and empirical uncertainty. Section IV offers a working definition of normative uncertainty, while Sections V, VI and VII present procedural justice and anticipatory governance approaches as means to account for this uncertainty. The concluding section offers some underlying worries of SAI research and climate change.

II.1.2 Solar Climate Engineering and Research Governance

Solar climate engineering[10] entails reducing incoming sunlight by increasing the earth's reflection level. One potentially highly effective way of doing this is by releasing aerosols into the stratosphere, an approach referred to as Stratospheric Aerosol Injection (SAI). On the basis of modeling[11], SAI is assumed to increase the reflection of incoming sunlight, which in turn leads to slowing down the warming of the planet, hence the oft-used umbrella term 'Solar Radiation Management' or 'Modification'[12] (Crutzen, 2006; Niemeier & Tilmes, 2017; Robock et al., 2009; Shepherd et al., 2009).[13]

However, the prospect of intentionally manipulating the climate is highly contentious, and passionately debated throughout the climate ethics literature. Opponents of researching the approach highlight how SAI may perpetuate unjust power structures, and represents the pinnacle of human hubris (Hamilton, 2014; Schneider, 2019). Others criticize how it is a form of 'moral hazard', wherein the possibility of climate engineering might distract from mitigation efforts

[10] Other common terms are: solar geoengineering, or just geoengineering/climate engineering. I use the term solar climate engineering, specifically focusing on Stratospheric Aerosol Injection.
[11] also extrapolating from empirical data and observation made during the Mount Pinatubo Eruption in the early 1990s.
[12] Whether we use the term modification or management is itself morally and politically interesting – do we control ('manage') or do we simply engage ('modify').
[13] Note that not all scholars agree with this taxonomy (Buck et al., 2014).

(Gardiner, 2009), or how its narrow focus on merely one aspect of the earth system is untenable as a solution to the complex challenge of climate change (Hulme, 2014). A main point of criticism that combines all the above, is that SAI carries incredible risks and uncertainties for the planet and its inhabitants, and is an immoral continuation of existing climate injustices.

To this central concern, proponents of research usually point out that the models so far show a decisive number of benefits of deployment over non-deployment in terms of risks, when compared to the current climate change trajectory (Crutzen, 2006; Keith, 2013; Keith and Irvine, 2016; Moreno-Cruz & Keith, 2013). Further, scholars highlight the necessity to engage in SAI research on prudential reasons. Due to its high impact on the climate and comparably low costs, climate economist Gernot Wagner for example touts the decision surrounding SAI as a situation of 'not *if*, but when' (Wagner, 2021). Importantly, SAI is not proposed by most of its proponents as an alternative to mitigation, but rather as a potential part of a diverse climate action portfolio, alongside thorough mitigation and adaption.

Moving beyond the controversy of whether research should be conducted in the first place, some scholars have proposed research governance frameworks. These seek to explore how an institutionalized SAI research program could be governed by, among other things, ethical considerations and principles (Gardiner & Fragnière, 2018; Rayner et al., 2013; Stilgoe et al., 2013; Tuana et al., 2012).

Mainly, these frameworks focus on how ethical considerations should play a central role throughout the research process[14] of SAI, emphasizing that '[. . .] ethical analysis be coupled with scientific analysis, and that funding agencies recognize the importance of such work through establishing funding programs for coupled ethical-scientific analyses of SRM'. (Tuana et al., 2012, p. 143). These integrated research governance frameworks seek to go beyond mere ethical analysis, but try to combine the science with the ethical, and socio-political issues at hand (e.g. supporting the political decision of research funding schemes and institutions).

An important aspect of these frameworks is the assessment of uncertainties arising throughout the research and potential deployment process. The science of SAI itself is riddled with

[14] Rayner et al. (2013) and Gardiner and Fragnière (2018) both blur the line between research and governance deployment, a point which Gardiner and Fragniere make explicit (2018, p. 145). While this merits attention in and of itself, it goes beyond the scope of this piece to take into consideration.

uncertainties (Kravitz & MacMartin, 2020). A comprehensive piece of scholarship on the intersection between governance and policy-making, the 2021 report on researching solar climate engineering from the National Academies of Sciences and Medicine (*NASEM*), states that

> A principal goal of any research program should be to better characterize and reduce scientific and societal uncertainties concerning the benefits and risks of SG [SAI] deployment (relative to global warming in the absence of SG [SAI]). (2021, p. 6)

It seems fairly obvious that any research effort should aim at reducing, or when possible even eliminating uncertainties related to the subject matter studied. However, there are important limitations to the degree of how much uncertainties can be reduced, as the report acknowledges.[15]

From an empirical perspective, the only way to deal with uncertainties is by engaging in more research, and learning as much as possible about SAI. This is no doubt necessary and should be a central role of any SAI research program. Yet, taking the potential irreducibility of some uncertainties seriously also highlights the fact that decisions surrounding SAI will have to be made under degrees of uncertainty.

The uncertainty of making decisions on SAI relates to both the societal and physical impacts of a potential research program. As the report claims, '[s]ome degree of uncertainty will be a persistent feature of SG technologies, especially at regional scales, as a direct result of uncertainties in the underlying climate model' (2021, p. 117). Further, the uncertainty does not only relate to empirically measurable facts. Uncertainty also arises with regards to the moral assessment of SAI research governance, for example how much (empirical) uncertainty is ethically acceptable when making decisions on SAI, or what kind of justice considerations should be given priority.

The following section analyses the decision-making space surrounding SAI research governance. Namely, decisions on SAI research governance arise in an ambiguous, multi-agential and multi-lateral decision-process that needs to reconcile diverse ethical considerations with incomplete empirical knowledge. These decisions take place under irreducible empirical as well as moral uncertainty, i.e. uncertainty that arises in value-pluralist contexts. This takes the current STS and

[15] 'Risks and uncertainties may be reduced, but not eliminated by research'. (National Academies of Sciences, Engineering, 2021, p. 116).

ethics of technology literature as a vantage point, exploring how this scholarship has so far attempted to deal with the phenomenon of uncertainty. I use the example of different renditions of the precautionary principle, which I claim are suitable to account for empirical uncertainty, but need to be adapted in order to deal with uncertainty arising on the basis of incompatible values.

II.1.3 Empirical Uncertainty and the Limits of Precaution

A main reason why risk and uncertainty are ever-present when thinking about new and emerging technologies, is due to the lack of foresight and prediction. There is no objective, incontestable data that a group of decision-makers could refer to in order to settle disagreements, or find a 'best solution'. The data and information policy-makers need to rely on in such situations is ambiguous, i.e. the relevant empirical considerations to reach a decision can all be reasonably contested.

The empirical facts surrounding SAI, and climate science more generally, are contentious and rely heavily on specific interpretations of data sets through models, and the construction of scenarios. Increasingly complex models aim to incorporate socio-economic developments, alongside climate responses (e.g. Integrated Assessment Models), and thereby become part of the decision-process. However, models are by definition not a complete representation of the real world, and while their inclusion is useful for making robust policy decisions, their inherent incompleteness and ambiguity gives rise to uncertainty for decision-makers.[16]

It should be noted that having to make decisions under risk and uncertainty in the policy realm surrounding technology and climate change is not a novelty. Philosophers and Science and Technology (STS) scholars alike have discussed the ethical and societal implications of modern technology, and the uncertainties they inevitably produce. Authors such as Ulrich Beck have argued that a central feature of modernity is its continuous production, imposition and resolution of risks imposed through and alleviated by technology (Beck, 1992). Similarly, Sheila Jasanoff points toward humility as a guiding principle when dealing with complex societal problems through technology and policy, in which '[u]ncertainty, indeterminacy, and ignorance are always present'. (Jasanoff, 2007, p. 33) The way new technologies shape society and the environment

[16] For a critical review of IAMs and their use in climate policy, see for example Frisch (2013).

are particularly pertinent in terms of their unintended consequences, and raises the question how humility and a new form of responsibility can account for these issues.

With the mainstream realization that the negative environmental and human health issues produced by neoliberal economic growth and exploitation paradigms are outpacing potential solutions, the so called 'Precautionary Principle' has gained prominence (Kriebel et al., 2001). The Precautionary Principle (PP) offers a way forward in situations of irreducible empirical uncertainties for environmentally impactful innovation and policy- decisions. Specifically, precaution highlights the importance of avoiding actions that could cause unacceptable harm, even in the absence of scientific proof that the action will cause said harm. In other words, when making a decision under empirical uncertainty, it is better to err on the side of caution.

There have been numerous attempts at synthesizing diverse accounts of the PP into a coherent, single definition (Gardiner, 2006; Steel, 2015). While certain core themes seem to form part of any PP, such as variations of Manson's *Activity – Effect – Remedy* tripod (Manson, 2002), I follow Lauren Hartzell-Nichols' argument that any PP needs to be appropriately contextualized and adapted to a given decision-situation (Hartzell- Nichols, 2017). In other words, there is not one Precautionary Principle, but rather a number of contextualized decision-making principles that highlight a contextualized form of precaution.

What all renditions of the PP do share, is that they seek to deal with uncertainties that are of scientific and empirical nature. These are uncertainties that arise in ambiguous situations wherein groups of decision-makers do not have access to the necessary information and therefore ought to better be 'safe than sorry'. Hartzell-Nichols, in her book on precaution and climate change, describes the uncertainties we face in climate policy poignantly:

> We are now quite certain *that* climate change is happening, but uncertainty abounds in our understanding of exactly *what* will happen, given different emissions and policy scenarios. We know climate change will be bad, but we do not know exactly *how* bad it will be *when* and *for whom*. We are therefore in a position where we must decide what to do about the risks climate change threatens in the face of a range of uncertainties. (2017, p. 27, italics from original)

All the italicized terms in the quote are descriptive pointers: The uncertainty is a matter of lacking knowledge and understanding.[17] This uncertainty that stems from a lack of knowledge is often referred to as 'empirical' uncertainty in decision-theoretical and policy literature (Walker et al., 2003). Marchau and colleagues for example describe making a decision under uncertainty as 'the gap between available knowledge and the knowledge decision makers would need in order to make the best policy choice'. (Marchau et al., 2019, p. 2) On their account, there is an 'uncertainty-gap' that represents the missing knowledge necessary to make a prudent choice. This uncertainty-gap is to be filled with empirical information, such as observable facts and data, hence the moniker 'empirical'.

Coming back to the PP, this empirical uncertainty is framed first and foremost as an obstacle for the decision-making process, as it blurs and contests available information. Examining the core dimensions of the PP, Per Sandin for example writes that 'uncertainty [. . .] expresses our (lack of) knowledge of [. . .] possible states of the world'. (Sandin, 1999, p. 892) However, this focus on empirical uncertainty runs the risk of overshadowing the equally impactful ethical aspects that ambiguous, value pluralist and multi-lateral decision-contexts need to account for. Importantly, empirical uncertainty impedes the unequivocal application of any given set of values or norms. If the impact of an action is shrouded in empirical uncertainty, the contextualization of moral judgments becomes uncertain as well.

Thus, beyond empirical uncertainty, the uncertainty related to decision-making on SAI research governance is also grounded in the moral realm, raising the question what kinds of values, norms, and ethical considerations motivate reasons for and against research in the first place. In other words, besides empirical uncertainty, decisions surrounding SAI research governance also give rise to *moral uncertainty* due to the empirical ambiguity and its multi-agential and multi-lateral decision-space.

II.1.4 Metaethical Moral Uncertainty and Practical Limitations

Originally, moral uncertainty is a term related to decision-theory and metaethics. However, I believe that there are two dimensions of moral uncertainty that need to be addressed for

[17] In an earlier piece, Hartzell-Nichols applies her version of the PP to the case of SAI, stating that research might be justifiable, but deployment certainly not (Hartzell-Nichols, 2012). The author reaches this conclusion by exploring whether deploying SAI can be seen as a precautionary measure to fend of catastrophic climate change.

clarification, specifically to contextualize it within SAI research governance. The first dimension can be referred to as meta-theoretical uncertainty, which is the dimension scholars currently mostly focus on. Scholarship dealing with this dimension of moral uncertainty seeks to find heuristic tools in order to overcome the uncertainty a single agent finds herself in, when trying to know what moral theory or concept is 'correct'. Scholars thus aim to answer the question how we should act if we do not know what the 'correct' moral theory is.

As an example, meta-theoretical uncertainty arises when one is uncertain about whether we should be vegetarians or not (MacAskill, 2016, p. 987).[18] Importantly, this meta-theoretical approach toward moral uncertainty treats the issue as a problem to be overcome through rational decision-making (MacAskill, 2016; MacAskill et al., 2020; Sepielli, 2014). Accordingly, in a recent book, MacAskill et al. (2020) try to deal with moral uncertainty by outlining an array of decision-theoretical tools, relying on the 'maximization of expected choiceworthiness' when comparing different theories. The decision of choosing a moral theory is relegated to a meta-theoretical equation, that produces a numerical value upon which the 'best' theory for a decision can be assessed.

Although MacAskill et *al's* approach is based on complex theoretical structures, their goal is to answer a practical question: How should a decision-maker act if they do not know which moral theory is correct? Their focus is explicitly on dealing with the uncertainty of moral questions as they relate to our individual decisions and the underlying reasons for these decisions, those reasons being moral theories. Should I tell a white lie or not, is translated into: Should I be a Kantian or a utilitarian?[19]

The meta-theoretical dimension of moral uncertainty assumes that moral uncertainty stems from disagreement about the ethical theories that underlie moral decisions. It is important to note here that we can take moral uncertainty to be an issue regardless of our metaethical stance, if one agrees that uncertainty occurs when making moral deliberations.[20] As long as it can be agreed upon that one *can reason* about morality in some shape or form, moral uncertainty is a relevant

---

[18] If one is a convinced vegetarian, there is of course no uncertainty about the matter. However, many struggle with the morality of eating animals, which makes this an illustrative example.
[19] It should be noted the term 'moral theory' is used here in a very broad sense. It also includes everyday moral principles and norms such as the golden rule ('Treat people the way you want to be treated'), for example.
[20] For an in-depth treatment of the challenges for non-cognitivists with regards to normative uncertainty, see (MacAskill et al., 2020, chap. 7).

phenomenon. A moral theory then does not necessarily entail any assumption of truth analogous to scientific truths, but it can answer questions such as what behavior and values are adequate, accepted, tolerated, etc.

While the meta-theoretical dimension of moral uncertainty is an important philosophical issue, it is not directly applicable to decision-making on SAI research governance. This is because the meta-theoretical dimension does not seek to account for the multi-lateral, value-pluralist and empirically ambiguous decision-context under which such decisions take place.

The uncertainty about the ethical aspects of a decision on SAI and research governance do not merely arise because decision-makers face theoretical struggles in coming to satisfying philosophical conclusions. Rather, moral uncertainty arises in its second dimension, which is the pragmatic necessity of making decisions based on contestable values, norms, or beliefs under empirically uncertain circumstances.[21]

SAI specifically produces moral uncertainty due to its global nature and uncertain impacts. It forces an international community to agree on the prioritization of a set of values that should govern the research of a technology with tangible global and intergenerational impacts, while those impacts are themselves uncertain. As with any decision that is being made in a value-pluralist context, some ethical considerations will take precedence over others. Some conceptions of justice, well-being, or autonomy will be given priority, especially given the fact that some ethical conceptions and their realization may be mutually exclusive.

This situation of multi-lateral value pluralism inevitably raises a specific kind of uncertainty among decision-makers, namely the practical issue of reconciling differentiating ethical positions that have no unequivocal reference point to settle the disagreement. We could of course claim that this is a similar situation to what philosophers and ethicists face when they try to account for

[21] I would like to thank an anonymous reviewer for pointing out that this second dimension of moral uncertainty raises an important question, namely whether it is simply a reframing of the struggle of coming to a shared decision under conditions of value pluralism. 'Moral' uncertainty would then fundamentally become a 'political' or pragmatic issue. My response to this is twofold. One, value pluralism (e.g. in the Rawlsian sense) is without a doubt part of what produces the uncertainty. Namely, I tacitly imply a metaethically value pluralist stance, which should entail a sense of humility when trying to make value- laden choices with others, who hold different values than me. Second, value pluralism alone is not what leads to uncertainty in SAI's case – it is crucially also the empirical uncertainty of 'the facts of the matter' and the political/pragmatist uncertainty that comes from the 'messiness' of having to reconcile and find compromise. Thus there is a combination of metaethical, empirical and political/pragmatic reasons for this dimension of moral uncertainty to arise.

meta-theoretical moral uncertainty. After all, MacAskill and colleagues' aim is to establish a method of reducing questions of moral uncertainty to an equation that solves for the highest 'expected choice-worthiness'.

However, this approach is not tenable when dealing with moral uncertainty in a global policy context. Policy-makers need to take into account not only coherent, philosophical reasoning, but (and arguably more importantly) the interests, claims, and rights of the constituents they represent. Whether and how to govern SAI research will ideally not be made by a single agent who finds herself struggling to combine a given ethical theory with ambiguous modeling results[22]. Rather, the decision-making process on SAI ought to entail broad participation by a number of experts and shareholders, with a specific focus on recognizing those most vulnerable to large-scale experimentation. These may all bring a number of differing ethical considerations, such as conceptions of justice and precaution to the table, which in turn complicate the decision-process.

At the same time, the contextualization of these ethical considerations can only be made under tacit assumptions of the kind of impacts the technology *might* have. However, those tacit assumptions can be reasonably challenged throughout, which calls the initial value-judgment they were based on into question. Given the empirical ambiguity and socio-political reality of climate policy decision-making, the meta-theoretical concept of moral uncertainty needs to be adapted to include moral uncertainty in policy-decisions.

II.1.5 Normative Uncertainty in SAI Decision-Making: A Working Definition

Based on the previous analysis of empirical and moral uncertainty, and the limitations of these concepts within the policy-space, we have now established a baseline to answer the question what kinds of uncertainties remain throughout the decision-process on SAI research. I argue that the uncertainties that remain are what I refer to as Normative Uncertainties (NU). Normative uncertainties are irreducible knowledge-constraints that arise in concrete, practical decision-situations. These decision-situations are multi-lateral and empirically ambiguous, wherein the decision-makers need to reconcile a plurality of values.

[22] I am – I think uncontroversially – assuming here a baseline of participation in the decision- making process as 'ideal'. More on this in the subsequent sections.

NU cannot be reduced to purely empirical or moral considerations. Rather, NU highlights the relationship between empirical facts and moral evaluation as central to any holistic ethical governance framework. While the 'facts' are invaluable, they seldom exist in a moral vacuum and their interpretation needs to be combined with an active reflection of the values and norms at play in the decision. At the same time, the ethical evaluation needs to be connected to the best interpretation of the current state of the world available, if it is to be both justifiable and feasible.

It is important to note that the concept of normative uncertainties I present here is not entirely novel. Especially with regards to climate policy, the taxonomy of normative uncertainties outlined by Taebi et al. (2020) needs to be mentioned. The authors introduce their notion of normative uncertainty as '[a] situation, in which there are different partially morally defensible but incompatible options or courses of action, or in which there is no fully morally defensible option'. (2020, p. 2) According to Taebi and colleagues, these uncertainties can be categorized in relation to the different reasons why they arise. Namely, they propose four different kinds of normative uncertainty: Epistemic, evolutionary, theoretical, & conceptual.

Epistemic normative uncertainty arises when there is insufficient empirical knowledge about a climate adaption proposal, which in turn raises ethical questions of evaluating uncertain impacts, and how to morally assess the imposition of risks and uncertainties. Evolutionary normative uncertainty accounts for the possibility of future generations holding different values than the decision-making generation. Since policy-decisions, especially regarding long-term, intergenerational infrastructure projects, are inherently guided by certain values, the possibility of those values no longer holding in the future needs to be taken into account. Finally, theoretical and conceptual uncertainty arise when different ethical conceptions or theories need to be reconciled throughout the decision- process. For example, the concept of 'justice' – what it consists of, how it could be achieved – is of central concern for debates surrounding political theory and philosophy that raises all kinds of conceptual and theoretical disagreements.

While the taxonomy of normative uncertainties makes up the majority of the article, Taebi et al. put forward a brief proposal for how such uncertainties could be dealt with on a decision-making level. Namely, they point toward a reflective decision-framework, inspired by John Rawls' Wide Reflective Equilibrium, and the emerging literature on adaptive planning (2020, p. 7). The next section engages with these possible solutions in more detail. Before moving on, however, it is necessary to position my proposal of NU in relation to Taebi et al's taxonomy.

There are a number of obvious overlaps between their approach and mine. Most importantly, I agree with the assumption that uncertainty is not merely the lack of measurable, empirical knowledge (what they refer to as *epistemic*), but also inextricably linked to moral considerations, such as value judgments and ethical background theories. Further, I concur with the notion that 'uncertainty in decision-making is not necessarily an undesired state of affairs that we should get rid of'. (2020, p. 4) While I have similarly described normative uncertainty as a knowledge-constraint, this does not entail that its ethical potential is purely negative. Uncertainty can also make room for previously marginalized voices and unthought moral considerations and scenarios. Uncertainty in policy-making can make space for flexibility and learning (Stirling, 2010).

With SAI research, irreducible uncertainty underscores the fact that there is a limit to how much certainty policy-makers can hope for when making a decision. The concept of NU I present here, can help with the decision-making process to identify knowledge- constraints throughout the decision's ethical evaluation. My focus on the procedural aspects of decision-making, as well as the decision-space in which it arises, contrasts it with Taebi, Kwakkel, and Kermisch's work. In this sense, I see my approach as a supplementary piece that contextualizes normative uncertainties beyond the climate adaption realm, and brings them into the discussion on new and emerging (climate) technologies, specifically SAI research governance.

This assessment of normative uncertainties then answers the first question of what kinds of uncertainties remain when having to make decisions on SAI research. With this backdrop of in mind, we can now turn to the second part of the question: How can we ethically account for NU?

II.1.6 Normative Uncertainty and Procedural Justice in Policy-Making

Given the empirically ambiguous, and value pluralist decision-context of SAI research, ethically reflected policy-makers are hard-pressed to find incontestable scenarios and undisputable moral values which could support their policy-decisions. Normative uncertainty challenges ultimate truths and unequivocal answers. However, if a clear assessment of what specific values should guide the decisions is impossible, we can turn to the procedure itself. In other words, instead of focusing on what kind of decision-outcomes are and are not justified, we can focus on the process of just decision-making itself.

In political philosophy, focusing on the just process of a decision is referred to as procedural justice. Procedural justice can be contrasted with the notion of substantive justice. Broadly and reductively speaking, the distinction between the two can be brought forth through their different answers to the question: What makes something just – its deliberative procedure, i.e. how the decision was reached, or the outcome of the decision? A purely proceduralist account of justice (Nozick, 1974)[23] would answer that what can be deemed just ultimately depends on whether a decision is based on a just process. Substantive justice approaches instead focus on the outcome of the procedure, e.g. whether certain goods were distributed in a just way.[24]

In line with procedural thinking, I propose that one way to account for the ambiguous and value-pluralist decision-space of SAI research, is through a process of robust participation. The normative uncertainty that arises necessitates a procedural, rather than a substantive approach for a simple reason. What the precise substance of justice is, what kind of distribution of well-being and harm among different regions for example, should matter, is what decision-makers – if a legitimate plurality of values and empirical ambiguity is taken into account – are uncertain about in the first place. Making substantive claims about who should bear which burden would leapfrog the procedure of justifiably reaching those substantive claims, and assume that there are unambiguous empirical facts and predictions available to substantiate those claims.

Following Axel Honneth and Nancy Fraser's seminal work on participation and recognition, the decision-procedure I propose relies on the normative assumption that a just procedure requires the 'parity of participation', i.e. that 'social arrangements [. . .] permit all (adult) members of society to interact with one another as peers'. (Fraser & Honneth, 2003, p. 36)[25] Translated into the realm of climate policy, where including the individual say of each and every affected member is virtually impossible, participatory parity can only be achieved through representation (Hourdequin, 2019, p. 35). The focus on representational participatory procedures allows to

[23] Robert Nozick himself maintains that whether or not something is just, depends on the history of its acquisition and just inheritance (through, e.g. contractual agreement).
[24] This can also refer to non-material, societal goods such as well-being, capabilities, or self- worth.
[25] Fraser & Honneth's notion of participatory parity also includes an 'objective', socio-material dimension. This dimension requires that all members of a society have the material, educational, societal, etc. means to actually meaningfully participate in a decision-process. (ibid 36ff).

account for both evolving empirical information as well as concomitant shifting values, norms, and judgments. Procedure provides flexibility as a means to handle the uncertainty.

As was pointed out, SAI research entails normative uncertainty due to the empirical uncertainty surrounding its outcome, as well as the political uncertainty of reconciling different values, based on the tacit metaethical assumption of some form of value pluralism. Participation and procedure can deal best with these uncertainties, given their focus on reflexivity and iteration, as well as their methodological agnosticism toward substantive ethical claims. However, the norm of participation in order to deal with normative uncertainty in climate policy has its own pitfalls of injustice, if it stays unqualified. For example, if we simply assume that justice is done if each country gets to have the same vote when a new climate policy is introduced, we fail to acknowledge the historical, socio-economic context of climate change. Unqualified, idealized participation can easily become a tool to perpetuate existing historical and structural injustices, by omitting the presence of underlying structural inequities.

In order to account for these pitfalls, some scholars have argued for the inclusion of recognition as a norm when it comes to climate justice and policy procedures. Recognition entails that the historical context of global power imbalances, based e.g. on patriarchal or racist-imperialist, neo-colonial politics, be taken as a background condition that needs to be accounted for in any decision-process. Justice as recognition within the procedural sphere then focuses on the need to '[understand] differences alongside protecting equal rights for all, especially given the uneven capacity to exercise and defend rights'. (Newell et al., 2021, p. 6; see also Whyte, 2011)

In terms of the ethical assessment of SAI research, recognition plays an important role in avoiding the reification of current injustices, through 'paternalistic attitudes that have historically characterized relationships between the powerful and the vulnerable', (unwittingly) ignoring the potential dangers SAI research might impose (Hourdequin, 2016, p. 43). At the same time, recognition is well-suited to account for normative uncertainty, due to its deliberative and iterative nature. Recognition can help identify the prevalence of normative uncertainty arising in those value-pluralist, multi-lateral contexts.

Including recognition into the decision-procedure on SAI research avoids the imposition of a culturally specific set of norms (e.g. Western-centric approaches) onto the global community. For example, proponents of SAI research, David Keith and Joshua Horton argue that '[SAI] appears to be the most effective and practicable option available to alleviate a range of near-term climate damages that are certain to hurt the global South most of all'. (2016, p. 89) However, this framing of SAI as a remedy for climate injustices overlooks the fact that climate justice, on the authors' own account of dealing with the 'global poor', would call for much more radical adaptation and mitigation measures, e.g. in the form of global reparations and redistribution.[26] In other words, this framing does not recognize the underlying structural injustices that underpin and exacerbate the harms of climate change.

The only way through which SAI deployment becomes the obvious candidate for redressing climate injustices, is if we assume a neo-liberal cost-benefit model, that accounts for climate harms in an a-historical context. These assumptions, in turn, could once again privilege those nations and communities already wielding most of the power, patronizing marginalized and vulnerable persons through the imposition of a specific set of norms in the decision-process. In other words, the failure to recognize and elevate the voices of those most impacted by potential SAI research leads to an exacerbation of injustices, not to their remedial. Such an approach would further seek to reduce the normative uncertainty, by implicitly claiming that a certain conception of justice, and a certain interpretation of the available data is to be prioritized over others.

At the same time, recognition must come from a place of respect and dignity if it is to account for justice in SAI decision-making (Hourdequin, 2019). In this vein, Olúfémi Táíwò and Shuchi Talati for example have cautioned that Western-based researchers, who call for a ban on SAI are themselves perpetuating colonial patterns by making choices on behalf of the global community (Táíwò & Talati, 2021). Similarly, the First Nations scholar and environmental ethicist Kyle Whyte points toward non-Indigenous environmentalists using Indigenous communities to argue for or against SAI research (Whyte & Buck, 2021). This form of 'present[ing] romantic accounts of Indigenous victimhood' (Whyte & Buck, 2021, p. 71) in order to make claims on behalf of Indigenous peoples, is an instance of misguided recognition in SAI research.

Thus recognition without respect and acknowledgment can easily be instrumentalized, especially in situations of normative uncertainty, where there are no unequivocal answers in terms of just

[26] I owe this reframing and critical understanding of Keith's and Horton's argument to an anonymous reviewer.

outcomes. For this reason, recognition in procedural justice needs to be accompanied by some clear guardrails, if it is to account for normative uncertainties in SAI research.

II.1.7 Qualifying Recognition and the Need for Normative Guardrails

Even an adapted, pragmatic concept of recognition cannot proceed without any ethical limitations, or otherwise normative uncertainty might open up space for radical relativism or nihilism. The fact that a group of decision-makers needs to account for different values and norms, which may at times be at odds with one another, does not yet incur normative uncertainty. Some values may be reasonable, based on conceptions of justice or well-being that can be justified, while others may not be. Normative uncertainty among decision-makers needs to be delimited by some shared fundamental guidelines, or principles, without being culturally imperialist or utterly relativist.

Regarding the international community, we can observe that there are some attempts of creating supposedly shared fundamental values, such as the Human Rights Charter, or the Sustainable Development Goals. Similarly, political philosophers have presented a variety of approaches that argue to capture all or parts of the human essence, on the basis of which they outline what individuals and communities should have access to[27]. What all these approaches have in common is to achieve justice – i.e. give an ideal outline of what justice could look like, and then seek to apply it to existing society.

In terms of uncertainty then, participants in a just deliberative process about SAI research need to consider what kinds of values should be part of a decision-process under NU, and what kind of values should not. For example, values based on imperialist efforts, racial and sexist discrimination and exploitation, or autocratic power motives should be excluded from a reasonably just decision-making process. In other words, a decision-space under NU needs to be delimited by ethical guardrails.

Guardrails throughout the decision-process would ensure two things. On the one hand, they would serve as limiting factors to what kinds of scenarios, futures, and data-interpretations are reasonable when dealing with ambiguity and value-pluralism. This is what Marion Hourdequin

[27] These approaches can be loosely understood as capability and function-based approaches, see for example (Nussbaum, 1990; Robeyns, 2006; Sen, 1987).

refers to as ‘basic respect’ (2019, p. 451), assuming the fundamental equality among persons and communities. Thus, SAI research programs need to ensure broad participation of all affected communities, appreciating the potential for different impacts. While the nature of the decision-process cannot account for all values, recognition would at the very least implore decision-makers to be clear about the values they choose to adhere to, and what kinds of moral considerations have been relegated or neglected. In this sense, recognition throughout the decision-procedure is a necessary condition for justice under NU in SAI research governance.

On the other hand, guardrails also free up the ethically justifiable space for argumentation, disagreement, and ultimately uncertainty. Accordingly, recognition does not only call for seeing other communities as equals, worthy of being included in the decision. Recognition also entails respecting important differences between communities (Hourdequin, 2019, p. 452) allowing for distinct cultural, societal, and moral consideration to play a role when evaluating a decision under normative uncertainty. Such respect requires to potentially accept norms and concerns foreign to one’s own value-system, creating a genuine space of open deliberation in a value-pluralist context. On this basis, guardrails create a free space of moral reflexion for policy-makers.

Note that these ethical guardrails should apply twofold. First, they should apply to the deliberative procedure itself, i.e. the process that seeks to explore SAI research. A possible instrument from deliberative democracy approaches are mini-publics. These are small, representative citizen groups (see for example Bächtiger et al., 2018) that could serve as an avenue to account for the procedural and recognitional elements of justice with regards to SAI research. Through the intentional engagement of representative parts of the public, deliberation on the research of SAI could become more just, rather than merely focusing on expert judgment and modeling outcomes. It could also lead to a more nuanced form of gathering public opinion (‘deliberative polls’), compared to traditional reflexive polling (Fishkin, 2009). Such a mini-public should ideally stay within such guardrails.

At the same time, mini-publics carry their own issues, by potentially upholding existing power structures unreflectively by virtue of the deliberative process that is still grounded in the unquestioned acceptance of political authority (e.g. Böker, 2017). In the case of SAI research, we can imagine that a mini-public might be used as a vehicle to further specific interests based on one-sided expert interpretations of data, which is especially relevant given the amounts of empirical uncertainty involved. Thus the guardrails need to also be applied to the creation of these deliberative spaces, so as to make space for critical reflection, especially if these deliberative spaces should be capable of designating exit-ramps and calling for the research to be stopped.[28]

The implementation of ethical guardrails sets up the playing field of reasonable ethical deliberation in the context of normative uncertainties when considering policy-decisions about SAI research governance. Delimiting such a space avoids the risk of simply justifying policies based on societal acceptance, highlighting the need for reflective ethical justifiability as well. (Taebi, 2017)[29] However, given the dynamic process of SAI research, another aspect needs to be taken into account when making decisions on SAI under NU, namely the processes' anticipatory and adaptive capacities.

II.1.8 Adaption, Anticipation, and Exit Ramps in SAI Research Governance

While a participatory and recognitional decision-procedure is able to account for the ethical challenges raised by normative uncertainty about SAI research, this approach alone does not yet outline how to practically implement such considerations. A policy-process is needed that allows for flexibility and revisions, and does not rely on static predictions of future outcomes.

Given the ambiguity under which SAI research would be planned, the process of making policies needs to be capable of revision and concomitant assessment. An ongoing research-process inevitably leads to new information, which in turn has to be included in the decision-making process on how to best proceed. SAI research governance is further a specific case, given its programmatic nature. Namely, an SAI research program would specifically aim to identify the technology's feasibility as a means to achieve a shared outcome (e.g. understanding why it's a bad idea[30], or whether there are just pathways toward deployment – these are all questions that a

---

[28] There remains a lot to be said on the construction of deliberative spaces, especially with regards to critical theory approaches. Thus the preceding section should merely serve as a means to exemplify how justice considerations could play a role in said construction.

[29] Taebi (2017) and Taebi et al. (2020) usually refer to the Rawlsian Wide Reflective Equilibrium in order to account for ethical justifiability. However, I believe that the inclusion of recognition should be front and center when it comes to SAI, given its global nature. Whether or not recognition can be a part of a practical WRE process, could in and of itself be an interesting research question.

[30] The motivation of having a research program to (also) prove its *infeasibility* tends to be overlooked, which I think should at least be a major question for SAI research.

deliberative process as described in the previous sections would need to answer). Given SAI's unique status of a 'non-existent /not-yet-existent' technology (Stilgoe, 2015), new information and insights are vital in defining its path, both for exploring its potential and realizing its flaws.

A suitable way to account for the normative uncertainty in the form of ambiguity is through adaptive planning. Adaptive planning entails that 'plans are designed from the outset to be adapted over time in response to how the future may actually unfold' (Kwakkel & Haasnoot, 2019, p. 358). Specifically, the concept of 'Dynamic Adaptive Policy Pathways' (Haasnoot et al., 2013), could serve as an appropriate 'policy architecture' (Kwakkel & Haasnoot, 2019, p. 361), i.e. the overarching conception of how a policy plan could look. Instead of assuming a central research question and static plan, it allows for iterative reassessment and reiterations of the initial policy pathway.

The iterative nature of Adaptive Planning further gives space for ethical reflection and robust engagement with all parties affected by the policy. Scholars from STS and the ethics of technology emphasize the importance of a preemptive and ongoing ethical assessment of new and emerging technologies through e.g. Technology Assessment approaches (Grunwald, 2015), techno-moral scenario-building (Boenink et al., 2010; Hofbauer, 2022; Keulartz et al., 2004; Swierstra et al., 2009) or Responsible Research and Innovation (Stilgoe et al., 2013). Among these approaches, the emerging concept of Anticipatory Governance could be particularly suitable for the ongoing ethical assessment of SAI research governance.

Anticipatory governance is 'a broad-based capacity extended through society that can act on a variety of inputs to manage emerging knowledge-based technologies while such management is still possible' (Guston, 2008, p. vi; see also Foley et al., 2018; Guston, 2014; Guston & Sarewitz, 2002). Anticipatory governance entails that decision-makers and publics are actively involved in the decision-process, and continuously evaluate, deliberate, and reflect upon the possibilities that a technology and its accompanying policy-decisions might produce. It thus allows for the reassessment of values, goals, and norms which guide the decision-process, while also being sensitive to the changes in the empirical understanding of the technology. By playing out 'what if . . . ?'-scenarios (Owen et al., 2013, p. 38) it is also specifically well-suited to deal with making decisions under normative uncertainty, exploring a plethora of scenarios and unexpected outcomes.

Adaptivity and anticipation should play a central role when it comes to SAI research governance, especially considering that SAI research may lead to the termination of strands of the research, or the entire program. This is necessary for a technology that is considered to be vulnerable to the development of unjust and harmful sociotechnical lock-ins. Such a lock-in situation of a SAI research program could unfold when institutionalized and vested interested (economic, political, etc.) create a research-context in which the original aim of learning more about the feasibility of a technology, turns into finding ways to create a pathway of eventual deployment. (Callies, 2019; Jamieson, 1996; McKinnon, 2019) This dynamic undermines ethically salient features of the decision-making process, by focusing on a- or immoral motives to continue the research.

In order to avoid such a lock-in, the NASEM's research governance proposal calls for the implementation of so-called exit-ramps, i.e. '[. . .] criteria and protocols for terminating research programs or areas' (National Academies of Sciences, Engineering, 2021, p. 9). Importantly, these ramps should be mechanisms that enable a swift termination of the research program. The qualification of these exit ramps also needs to be understood as an iterative process in and of itself, that is comprised both of recognitional deliberation and participation, as well as anticipatory and adaptive planning. Importantly, while some exit ramps may serve as guardrails themselves, by declaring certain research outcomes as morally unacceptable (e.g. the risk of catastrophic outcomes, see Hartzell-Nichols, 2012), others may be adapted throughout the research process. Given the possibility of changing values and newly found information, exit ramps must be both robust in their application and dynamic in their assessment.

II.1.9 Concluding Thoughts on SAI Research

Throughout this paper I sought out to explore what kinds of uncertainty will persists during the decision-making process on SAI research governance, and how these uncertainties affect the ethical assessment of such a potential governance project. I argued that current approaches tend to reduce uncertainty to either empirical or metaethical issues, which overlook the moral and practical dimensions of decision-making under what I have called normative uncertainty. Normative uncertainty, I argued, is a situation of ambiguous information and multi-lateral value pluralism that lacks unequivocal references in order to reach a decision. In order to deal with this normative uncertainty, I proposed the inclusion of recognitional, participatory justice considerations as well as adaptive and anticipatory governance methodologies.

While SAI research may produce invaluable insights on climate change beyond the technology itself, it must be stressed that even if feasible, SAI cannot be a solution to the problem of anthropogenic climate change. Importantly, any kind of SAI deployment cannot account for the underlying structural injustices that make climate change so harmful for so many in the first place, such as colonial land grabbing and displacement, climate racism, and the environmental exploitation for the monetary gains of a wealthy few. While SAI does not and cannot account for these injustices, a research program that seeks to reflect on the ethical assessment of such a technology should also be a space to reflect on the question *how it has come to this*.

III. Chapter 2: Putting Cats Back Into Bags. Disruption, Reversibility, & Lock-in In Solar Climate Engineering

This paper critically examines the normative guidance as well as practical feasibility of reversibility as an ethical criterion for New and Emerging Technologies, with a focus on solar climate engineering through Stratospheric Aerosol Injection. Specifically, concepts such as lock-in, control, and disruption, all relate to some degree to the idea that reversing unwanted technological impacts is both morally desirable and prudent given the risks and uncertainties NESTs introduce. At the same time, the exact implications of what should be reversible remains unclear, whether it be societal institutions, value systems or collective worldviews, physical and non-physical infrastructure, etc. The claim is that while a useful tool for the investigation of new technologies, reversibility needs to be strictly qualified and contextualized, if it is to serve as a criterion for ethical and practical guidance.

III.2.1 Introduction

New and emerging technologies (henceforth *NESTs*) bring a lot of opportunities, as well as numerous risks and uncertainties. This paper explores solar climate engineering through Stratospheric Aerosol Injection (*SAI*) as a case study of a potentially highly disruptive NEST. SAI is considered by some as a possible additional measure to reduce the pace of global warming, by releasing aerosols into the stratosphere. This would scatter some of the incoming sunlight, and consequently cool the planet (Crutzen 2006; Keith 2013). Throughout the paper I focus on SAI as imagined through aircrafts that reach the stratosphere and continuously release some reflecting agent, the most commonly referred to form of solar climate engineering (Keith 2013; Robock 2008; Shepherd et al 2009; NASEM 2021) Given the uphill battle humanity faces against climate change, SAI is a prevalent, albeit polarizing topic within the climate ethics literature (Preston 2013; Pamplany, Gordijn, and Brereton 2020).

The premise of this piece is that the ethical acceptability of SAI research and potential deployment depends on three central concepts, namely control, disruption, and reversibility, with reversibility being the central focus. SAI serves as an insightful case study for the exploration of these concepts for two reasons that taken together, produce an interesting tension. First, reversibility serves a normative function. Whether or not SAI research and deployment is justifiable, will greatly depend on the degree of reversibility of its impacts, given the many uncertainties Second, SAI is a not-(yet)-existent technology, rather than a technology waiting to be deployed. SAI and its research put pressure on reversibility's normative function, raising the question what exactly ought to be reversible and what can be reversible in the first place. In a nutshell, while it seems prima facie prudent to ensure SAI's impacts to be reversible, a closer look shows that it is unclear what exactly ought to be reversible.

The paper is structured around investigating these two aspects of reversibility in SAI and its research. SAI and its research have complex, interdependent and cross-boundary, sociotechnical impacts. These sociotechnical impacts are comprised of, roughly speaking, a society's individual and collective value-system, its institutions, be they political, economic, or other, and its physical as well as non-physical infrastructure. The kind of reversibility that I argue is most crucial to SAI research then, is also the most complex and intractable, given both the technology's inherent uncertainty and its non-existent status. The main contribution of this piece is to the ethics of technology scholarship, as well as science, technology, and society studies. Contextualizing SAI research within these strands of research, gives a new angle in the ethical assessment of not only solar climate engineering, but other potentially disruptive emerging technologies as well.

The first section of the paper gives an overview of reversibility and how it relates to technological development, as well as the role the concept has played within the ethics of technology. Section two ties reversibility to the concept of control and lock-in. On that basis, section three builds on discussions of sociotechnical structuration to explore how the research and deployment of a technology can reinforce existing societal structures. Section four investigates how technology can be disruptive, having the opposite effect of entrenchment, forcing society to adapt its structures or create new ones. This section also deals with the question to what degree a technology has to be deployable, in order to have an impact through the example of SAI. The final revisits the normative relevance of reversibility for NESTs.

III.2.2 Reversibility, Control, and Lock-In

The concept of reversibility has found applications in the philosophy of science, economic theory, and environmental ethics and economy (Humphrey 2001; Verbruggen 2013; Pols and Romijn 2017). From an innovation analysis perspective, Merkerk and van Lente have argued that emerging irreversibilities create an innovation path, reifying a set of technological developments and shutting out others (van Merkerk and van Lente 2005). This form of irreversibility reduces the complexity surrounding the development of new technologies. These emerging irreversibilities create constraints and depth. The situation that constitutes the irreversibility precludes certain options, but increased predictability enables more specified development. Through the creation of expectations and agenda building (e.g. creating roadmaps, institutional and financial support, conferences, public information, etc.), specific visions of the future are performed, and the technology takes shape through the research and design process, partially guided by irreversibility. Accordingly, a certain amount of irreversibility in the form of a "shallow lock-in" is not only inevitable, but necessary for technological development (Shackley and Thompson 2012, 119).

However, the lock-in necessary for technological development is also a central point of criticism for NESTs in general, and SAI in particular. Worries of technological lock-in are usually worries about continuing a certain research trajectory for misguided (e.g. economic or political) reasons, while also foreclosing possible alternative futures. Akin to a slippery slope, lock-in represents a loss of societal control over a technology's development: "The thought is that research sits at the top of a slippery slope, at the bottom of which awaits the full-scale deployment of a morally objectionable technology." (Callies 2018, 676) In the case of SAI, given the socio-political context of its research, critics claim that there is a high risk of such a lock-in (Hamilton 2013, 2014; Cairns 2014; McKinnon 2019), and a loss of societal control over the technology's trajectory.

Focusing on the controllability of technological development is the flipside of lock-in, and gives a hint towards the normative relevance of maintaining control and avoiding lock-ins. Namely, the focus for ethical technology assessment should be the "[m]aintenance of the social control of technology and the prevention or amelioration of negative impacts of technological development" (Cairns 2014, 651). In short, any assessment of NESTs ought to focus on allowing for control and openness for change in the technology's trajectory. Reversibility in the context of lock-in is accordingly a salient criterion for the ethical assessment of NESTs and their innovation. Specifically, a technology should be assessed by the amount of societal control it allows for, before, during and after implementation. The focus on societal control is based on the assumption that developing and introducing a new technology, with all its accompanying risks and uncertainties, should allow for the possibility to steer the technology's development and potential subsequent deployment (Genus and Stirling 2018). Maintaining this kind of societal control over the development of a technology lies at the heart of the oft-cited "Control Dilemma" or "Collingridge Dilemma" (Collingridge 1980). The Control Dilemma is constituted by an imbalance of controllability between a technology's development phase and its deployment. NESTs are malleable in the initial phases of research and design, while how the technology will be used in a societal context is not yet well known. Once the technology's use becomes apparent and entrenched in the societal context, and thus better understood, it becomes more difficult to change.

Given the impossibility of anticipating a NESTs impacts, maintaining societal control over them is paramount. This entails keeping open a possible future where the technology is no longer used, or its use is changed. Relatedly, reversibility and the avoidance of lock-in through technology is especially important in terms of unintended consequences, and the potential for a loss of control over the technology as NESTs produce "deep uncertainties", i.e. inherently unforeseeable consequences (Marchau et al. 2019). For example, a new technology can create a kind of dependency that forces their continuous use throughout the future. They might also produce new problems through their continuous use. But the problems they cause might only arise after they have been fully entrenched, making the issues chronic and the only practicable treatment symptomatic. The development and planning of cities to account for individual transportation via cars for example, has led to very specific kinds of urban environments that are difficult to reverse. Physical structures, market mechanisms, ways of life and politics, as well as institutions have been built around the concept of individual transportation through the automobile. This entrenches the perceived necessity of cars, and crowds out potential alternatives. Similarly, the reliance on a fossil-fuel based energy system has created a deeply entrenched dependency that is extremely resistant to any kind of change (Dryzek and Pickering 2019; Shackley and Green 2007), which is a major challenge for the curbing of emissions. The potential loss of societal control over any NEST is therefore clearly ethically relevant. Whether or not a technology's trajectory and impact can be controlled by being reversed is particularly important when trying to assess the impact of NESTs. How this impact might shape society and our environment is further directly relevant when it comes to evaluating whether the technology should be researched or deployed at all. Reversibility is therefore a necessary part of maintaining societal control over NESTs the innovation process, a central challenge for SAI research.

III.2.3 SAI, Structuration, and Entrenchment

SAI's uncertain societal and environmental impacts are some of the reasons why the technology serves as an insightful case study for the exploration of reversibility. The research that is being done on SAI in the form of modelling, highlights the fact that the approach inevitably comes with some irresolvable uncertainty, especially on the regional level (Kravitz and MacMartin 2020). How and to what degree SAI deployment would impact precipitation patterns, cloud formation, and regional weather in general is virtually impossible to predict.

While the climatic impacts are an important worry for the development of SAI, contemplating and researching the technology could already reinforce existing values, norms, institutions, etc. Scholars like Naomi Klein have explored how norms, and world views that guide many societies and the global market today, such as consumerism, neo-colonialism, and unbridled capitalism, have led to climate change, and the fossil fuel dependency humanity currently faces (Klein 2015). A society that relies on researching "techno-fixes" for complex societal problems (Huesemann and Huesemann 2011), relies on the alienation from nature, and the exploitation of marginalized communities, while inhibiting the societal engagement of oppressed groups, such as indigenous peoples (Schneider 2019; Dryzek and Pickering 2019; Whyte 2018). Accordingly, critics of SAI frame the very idea to be representative of a kind of imperialist worldview that requires centralized and authoritarian power structures to work in the first place. Critics further argue that the institutionalisation of SAI research through funding and political support will inevitably lead to some kind of deployment, given the politico-economic interests involved (Hamilton 2013; 2014). The research and subsequent deployment would not be motivated by justifiable scientific and moral reasons, but by the political and economic interests of a wealthy elite. A research program would thus lead to a morally objectionable lock-in (Cairns 2014; McKinnon 2019; cf. Hale 2012).

Drawing on this criticism, some have a more generous outlook on researching SAI. Eric Winsberg for example argues that even if research would lead to the conclusion that SAI is not feasible, there might still be epistemic value in building a research program (Winsberg 2021). Further, Heyward and Rayner observe a "curious asymmetry" when it comes to how SAI, as a potentially global intervention is being treated, compared to radical mitigation measures (Heyward and Rayner 2013). Both require vast oversight and thorough political intervention, but only the former is considered potentially undemocratic for it.

These arguments warning against the very idea of SAI, presume a kind of sociotechnical structuration through NESTs that makes reversing course in the future difficult, if not impossible. Namely, technologies can have an entrenching effect on existing values, norms, institutions, physical structures, etc. This is because the repeated societal use and an ultimate reliance on a technology makes changing it increasingly difficult. The irreversibility of the change through technology is reinforced through a process of structuration (Bergen 2016a, 2016b; Orlikowski 1992; Giddens 1986) that supports and entrenches the development in the form of discursive, material, and institutionalized reproduction. Structuration follows the line of thinking of the politics and philosophy of technology more generally, in that technology and its innovation are inherently value-laden and political processes. Technology and innovation are imprints of values through concrete, metal, and plastics, with the built environment enabling specific behaviours and restricting others (Brey 2018), representing its own "politics" (Winner 1980). Again, planning cities with individual cars in mind serves as an illustrative example here. With the technological achievement of the automobile, personal mobility through the use of cars was considered a central claim to freedom (in the US for example, qua the "American Dream"). While this value might appear slightly anachronistic today, in many cities, not owning a car converts you into a more vulnerable pedestrian or cyclist. The car is considered to be the "normal" means of transportation.

This consideration is actualized in the physical world through parking lots, roads, traffic lights, and sidewalks. Importantly, certain kinds of NEST can reinforce, rather than undermine this development. For example, electric vehicles do not undermine the value and expectation of individual transportation, nor the institutions and politics surrounding the automobile. They just seek to 'fix' the unintended consequences, i.e. pollution and CO2 emissions, by tweaking the original technology. This in turn, reinforces existing infrastructures, such as car cities and some form of publicly available charging stations, parking lots, etc. Accordingly, the introduction of a new technology could lead to the conservation of current societal structures, making any form of change more difficult, and entrenching the status quo.

Reversing these structures, and the values they express, would require the reversal of the built environment through physical activity, such as constructing new infrastructure or tearing down old ones. This is because the physical infrastructure serves as a continuous enabler of individual mobility through cars, which in turn affects the personal, professional, and institutional expectations between human beings. However, a central lesson from the philosophy of technology is that the physical reality of technological artefacts does not arise in a conceptual vacuum. Rather, technology always already frames the environment in a certain way (Heidegger 1977), and not unlike morality, mediates the way we interact and see the world (Latour 2002; Verbeek 2011). The way technology mediates, and thus (re)-shapes our aims is central to the question what kinds of technological innovation we ought to seek, and what kind of reversibility beyond physical structures is prudent. The question then is whether SAI and its research might foster the emergence of new values, norms, institutions, and ways of seeing the world, and whether it is feasible to seek for the reversibility of such an impact.

III.2.4 SAI and Societal Disruption

Apart from entrenching existing norms, values, and institutions, NESTs can also be socially disruptive. The idea of social disruption is based on "… how technologies can disrupt social relations, institutions, epistemic paradigms, foundational concepts, values, and the very nature of human cognition and experience." (Hopster 2021, 1) NESTs thus potentially undermine societal norms, practices and institutions, or put into question ethical values and concepts. Importantly, if we assume that NESTs can be socially disruptive, a prudent measure would be to ensure that their impacts can be reversed.[31] If a technology is disruptive, it creates new norms and values or challenges existing ones. It might serve as the basis for new world views and cosmologies, tied to institutional and political changes, or undermine apparently stable societal structures. New technologies' potential to disrupt and undermine or challenge shared societal imaginaries, values, and norms, has received specific attention in the literature on techno-moral change (Keulartz et al. 2004; Swierstra, Stemerding, and Boenink 2009; Boenink, Swierstra, and Stemerding 2010).

[31] Disruption and entrenchment often occur simultaneously, albeit on different societal levels. What is considered disruptive for one group in society, might reify the practices of another, or have no effect at all. Disruption through technology is contextual (see also Hopster 2021 and Socially Sustaining Technologies).

A salient instance of such a techno-moral change is the liberating push for sexual freedom in the 1960s, coinciding with the introduction of the contraceptive pill (Baker 2019).

Disruption occurs as a new technology's introduction tips the scale of societal value disagreements in favour of one value, or another set of values. This can undermine prevalent norms or introduce new ones. For example, the introduction of the contraceptive pill greatly reduced the practical concern of unwanted pregnancies, which further undermined the patriarchal and sexist values of strict (female) chastity and purity. This was a consequence of the pill allowing for women to obtain some control over their own body, deciding if and whether to get pregnant on their own terms.[32] Disruption through technology can further manifest itself in the form of moral uncertainty. Such uncertainty arises when the way we perceive the world and our moral relationship to our surrounding, is re-negotiated through the introduction of a new technology. Prenatal diagnostics is one such example, where the use of a technology forced new responsibility onto doctors, health-care workers, and parents (Verbeek 2011). Similarly the question of assigning responsibility in the case of accidents involving human life and automated vehicles raises pertinent issues in terms of culpability (cf. Nyholm 2018).

Moral uncertainties arise for different reasons, but they are all based on the problem of not knowing how to ethically assess a given situation, or how societal norms and values develop (MacAskill, Bykvist, and Ord 2020; Taebi, Kwakkel, and Kermisch 2020; Hofbauer 2023). Technology can undermine existing moral concepts, which makes deriving normative conclusion from them difficult. The introduction of the mechanical ventilator for example, kept patients without a brain function "alive". In order to understand this new situation that the patients found themselves in, and make moral decisions about how to handle them, the concept of "braindeath" became necessary. (Nickel 2020, 20; Baker 2019) Besides undermining existing values and norms, and creating normative uncertainty, NESTs can further give rise to new moral considerations and values. An example of this is the increasing worry of data protection and privacy over the last decade, coinciding with the implementation of big data technologies such as social media, streaming platforms, online shopping, and many more (Véliz 2021).

---

[32]The pill did not subvert all societal norms, which becomes especially clear in terms of maternal expectations in a patriarchal society. Sexist norms did not vanish with the introduction of the pill, highlighting the limitations of technological disruption.

Similarly, SAI challenges current norms and pushes to favour certain values over others. For example, SAI challenges the notion of “pristine” nature, since its effects are global by design, and thus leave little room for nature that is untouched by humanity. This disruption could change the societal relationship with the environment, and our direct relationship with other human beings, once the arteficiation of the climate is concluded (Hofbauer 2022). Researching and deploying SAI might push a hyper eco-modernist view of nature, wherein the natural becomes nothing more than a resource for human ends (Latour 2017; Neyrat 2020). At the same time, this global impact of SAI also lends itself to the idea that sustainability in the age of the Anthropocene can only be achieved through active intervention (Steffen et al. 2011). The conception of nature being vulnerable to human destruction while at the same time requiring human intervention to persevere lies at the heart of this perspective, also referred to as earth stewardship (Kolbert 2021).

If new technologies lead to this shift in societal structures, being able to reverse them is akin to ridding society of newly emerged norms and values. If one were to call for SAI’s societal effects to be reversible, one must then refer to its disruptive potential. Effectively this means that the new values that the introduction of SAI produces (active stewardship of the earth system), and the old ones it undermines (pristine nature), should potentially be abandonable or retrievable, respectively. Presumably, reversibility would then entail regaining certain values that were abandoned after the introduction of the technology. It seems that such a general call for irreversibility is hardly feasible. However, a more specific sense of narrow reversibility should still be achievable. This is what scholars reflecting on ethical SAI research advocate for, when they call to maintain societal control over the technology’s development (Rayner et al. 2013; National Academies of Sciences 2021).

It is important to note that such a form of societal control does not guarantee the reversal of changed norms or values. The very process of research could already shape our conception of nature to such a degree that it reinforces one valuation of nature over another (Midgley 2006). If these impacts of NESTs ought to be reversible, such reversibility would refer to a sort of openness for the future, leaving future societies with the possibility to choose different paths, and to realize values and norms that would otherwise be unreachable.

III.2.5 Forms of Disruption, & The Question of Speculative Ethics

NEST's societal impacts ought to be reversible, since their research and deployment might be harmful. However, not all NESTs have the same kind of impact, and might be disruptive and irreversible on different levels. To illustrate this, we can consider two levels of disruption, one that affects existing structures and processes and another that introduces new ones. I refer to the first level as *adaptive* disruption, and the second level as *novel* disruption. I argue that NESTs leading to adaptive disruption may be more easily accepted by society, while their impacts are potentially less profound and more reversible. Novel disruptions are potentially more difficult to accept, but become harder to reverse once they have gained traction.

Adaptive disruption can be explained through the example of autonomous vehicles. Society puts laws, infrastructure, market-mechanisms, institutions, etc. in place, which serve as a framework for individual transportation via car. These structures are the framework within which motorized individual transportation takes place. The introduction of the self-driving car directly impacts this structure by changing certain elements of the previously existing framework. Specifically, driverless vehicles do not have a driver that could be held responsible, is marketed to, needs insurance, etc. The existent structure needs to now adapt to this impact, which means the impact necessitates a reaction from the original institutions, laws, infrastructure, etc. in order to cope with the new technology.

Adaptive disruption occurs if there is a structure in place, both on a conceptual (society knows what a car is, what its function is, why it matters, imbues it with values) and an institutional, legal, infrastructure, and political level. The introduction of a new technology then necessitates this structure to be adapted to a different practice, in this case the practice of co-existing, using or being impacted by, driver-less cars. However, the introduction of the technology is not a fundamental disruption. People already accept the value of individual transportation, and how it could be handled. Therefore, the introduction of autonomous vehicles, while potentially disruptive on some levels, could ultimately be socially acceptable. At the same time, their introduction is certainly reversible, given that they rely on existing structures.[33]

The case of solar climate engineering through SAI serves as an illustrative example of a novel disruption. The intention of SAI is unprecedented in that it has no previous structures that seek to accommodate a similar goal. There are no structures in place that would guide the intentional manipulation of the climate through technological means. Certain types of infrastructure exist that could facilitate its introduction, such as airplane industries, airports, chemical plants, international treaties and governance bodies, and others. But the purpose of intentional climate manipulation poses entirely new questions, and challenges societal norms and values fundamentally. Humanity has no clear idea of what it would mean to 'manage' or 'manipulate' solar radiation, with the purpose of reducing global warming. Consequently solar climate engineering through SAI might be a case of technology shaping societal goals in a profound manner. Bruno Latour discusses this quality of technology in the following way:

> If we fail to recognize how much the use of a technique, however simple, has displaced, translated, modified or inflected our initial intention, it is simply because we have *changed the end in changing the means,* and because, through a slipping of the will, we have begun to wish something quite else from what we at first desired." (Latour and Venn 2002, 252, emphasis from original)

The potential for SAI to cause society to "wish something quite else from what we first desired" is clearly a central worry for those opposed to SAI research. What may initially be framed as a means to combat climate change ("what we first desired"), suddenly turns into ambitions of global climatic control, unlimited natural capital, or authoritarian power fantasies.

The non-existence of SAI as a tangible technology raises another important concern in the form of "speculative ethics" (Nordmann 2007). Speculative ethics is a critical term, introduced by Alfred Nordmann to problematize the assumptions made by ethicist and philosophers engaging in imagined far-away futures, wherein technologies are framed to develop in a certain way. Specifically, the criticism focuses on what Nordmann calls the "foreshortening of the conditional" or the "if and then"-syndrome.

What Nordmann points out here is that ethicists tend to underestimate the "if" part of technology assessment. Editing the human genome to achieve immortality, space travel outside our solar system, android robots, etc. are all examples of things that *could* occur through technology and innovation (the *if* part), so we should start thinking about their ethical implications and how to handle them (the *and then* part). The *if* gets treated by speculative ethics as if it would occur anyways, and as if questioning ever-increasing technological progress is basically naïve.

These are important criticisms and reflections on the innovation-myth that drives modern societies. While certain aspects of the “if and then”-syndrome can certainly be found in the climate engineering debate, the critique needs to be contextualized within the circumstances of SAI research governance. The debate on how to deal with the “vanguard vision” of SAI (Hilgartner 2015) is centrally a debate on to what degree humanity should invest its limited resources into such an approach. A focus on research is precisely because the technology does not exist (yet): Should it exist? *Could* it even exist in the first place?[34]

Asking whether SAI should and could exist as a deployable technology can be done without having to “foreshorten the conditional” (Nordmann 2007, 34), i.e. without having to act as if it already existed or will certainly exist. Whether and to what degree humanity should engage in the research of SAI raises critical questions of justice and ethics, for example (e.g. Preston 2016; Pamplany, Gordijn, and Brereton 2020) that give insights on issues of climate justice and global equity beyond the issue of climate engineering. Further, SAI is already having real-world effects. Arguably, the call for a research ban (Biermann et al. 2022) and criticisms of SAI’s potential to lead to mitigation deterrence (McLaren 2016) highlight the fact that the idea alone has a socio-political effect. That effect is in and of itself morally relevant, notwithstanding the technology’s feasibility. And whether that effect is in any way reversible is clearly relevant. Namely, can we learn something from SAI research, beyond whether the technology works?[35]

With this backdrop in mind, I believe SAI and its research can be reasonably investigated by exploring its potential sociotechnical impacts, without relying on exact predictions. This brings us back to adaptive and novel disruption, and their respective reversibilities. Institutional and societal structures often move slowly when adapting to a new technology. For this reason I tentatively argue that adaptive disruption is more reversible than novel disruption, and more easily accepted. The slow pace allows for some time to debate and deliberate the new technology. However, it is also easier to have a slow-moving acceptance of the technology. Resistance to the technology can more easily be whittled down, so to speak.

Novel disruption potentially faces a lot more societal pushback, while its impacts potentially end up being less reversible. The example of solar climate engineering shows that the strangeness, coupled with the apparent hubris and obvious risks of SAI, is difficult to accept. This could be seen as a result of the technology's imagined traits of being inherently unpredictable or controllable, even for those who currently wield power. If we assume the dominant global order to be one of "imperial market globalism" (Steger 2007), we find that SAI is not immediately the entrenching force that some of those wielding power might hope it to be, or critics claim it would be. As Jeremy Baskin writes in his discussion on the framings of SAI:

> On the one hand SGE [SAI] is a response emanating from the centres of global power, rooted in standard practices and epistemologies of science, and purporting to be a solution to the climate change problem. […] *And yet it is simultaneously a subversive story which undermines a great deal in the dominant accounts of modernity, of science and of the human-nature binary*. (Baskin 2019, 222, my emphasis)

It is also a lot more difficult to have a process of normalization and acceptance of the technology, since slow-paced gradual introduction is not possible in the case of SAI. We either deploy SAI, or we do not, and its effects are necessarily global. The distinction between adaptive and novel disruption gives an outline of how NESTs in general and SAI specifically affect societal structures. On the one hand, adaptive disruption occurs when existing structures are affected, and need to adapt to a new technology. On the other hand, novel disruption occurs when there are no structures in place which could fit the bill of accommodating, enabling, or controlling a new technology, given the technology's fundamental novelty.

III.2.6 Revisiting Reversibility as an Ethical Criterion

The reason many scholars invoke reversibility in the first place (even if often merely implicit within the ethics of technology literature), is to avoid lock-ins of unfavourable or harmful technology-induced pathways. In the case of SAI, a central piece of criticism is that the technology could lock us humanity and the planet into a potentially unjust and dystopian future. Importantly, SAI research could create a form of entrenchment that reinforces existing, unjust and unsustainable societal structure. Instead of eliminating society's carbon dependency, and the underlying power structures that perpetuate such a dependency, we simply rely on an incredibly risky and potentially unjust techno-fix. Crucially, despite realizing that the technology has unfavourable impacts, the process can no longer be aborted due to its inherent lock-in through the looming risk of a termination shock. In short, the worry is that SAI is an essentially irreversible technology once deployed, and its research might already set humanity on track to conserve the values and norms that caused catastrophic climate change in the first place.

On this basis, I argue that the ethical relevance of reversibility is in need of precise qualification. This stems from two main reasons, one moral the other practical. First, the moral value of reversibility is not clear, unless it is tied to some norm or value that should be retrievable. Thus maintaining societal control over technological developments, is crucial in order to ensure that the development adheres at least in some form to societal values. New technologies ought to allow for flexibility and adaptivity with regards to the emergence of new societal values.

However, reversibility spelled out in terms of keeping all options open, cannot itself serve as a moral criterion. Concretely, there seems to be no moral reason to always enable future societal pathways that might produce great harm. Enshrining inalienable human rights into constitutions, for example, is a way of ensuring the conservation of some moral beliefs on a societal level. Therefore, there needs to be a balance between allowing for flexibility in the decision-making process and the reversibility of societal harms, versus potential societal progress. As a case in point, the introduction of the contraceptive pill played at least a partial role in the emancipation of female sexuality, and sexual liberation more generally from patriarchal gender norms. From a feminist and equity standpoint, it seems that such changes *should not* be reversible. Similarly, if the introduction of SAI could have a positive impact (whatever that may entail) on how society values nature, it seems that such a change ought to not be reversible. However, this does not warrant sacrificing any and all reversibility for the sake of perceived potential progress, however defined. Reversibility needs to be infused with certain values in order to be normatively action guiding.

Second, specifically the sociotechnical reversibility is in need of refinement in order to be practically action-guiding and feasible. Whether or not it is feasible to have the option of undoing physical infrastructure and impacts of technology, is dependent on political will and economic reasoning, which are both too broad of issues to be done justice here. However, the physical infrastructure is fundamentally linked to a technology's societal impacts. The reversal of those societal impacts seem to be a less straight forward criterion for technological development. Societal values and norms can develop in a plethora of ways, and manifest themselves through numerous outlets. While technology plays an important role in those developments, the genesis of societal value change in relation to technology is by no means clear.

The dynamics of technological disruption, and where to pinpoint its origins, are contested and impossible to pin down to a single factor. Without existing, SAI already has an impact beyond purely academic debates, such as the plan to invest into solar climate engineering research from the United States government.[36] Consequently, one is faced with the issue *where* precisely the reversibility should manifest, and when a technology starts having an impact. It could be on an institutional level, within a legal framework, through education, or through grass roots societal movements – whichever realm one looks at, it quickly becomes clear that a singular instance of a disruptive force cannot be identified.

To sum up, reversibility can be an insightful tool to better understand the risks new and emerging technologies entails. Technological development is often shrouded in uncertainty with regards to the ethical issues it brings. Reversibility can be a conceptual lens through which potential moral worries about these developments can be analysed. Importantly, reversibility can serve as a first attempt to understand the impact that NESTs such as SAI could have on a societal level, be it through further entrenchment of unjust and unsustainable societal structures, or the disruption of existing structures. However, the concept lacks normative guidance as well as practical feasibility, if it is applied without a clear context and thorough qualifications of *what* and *how* something should be reversible. Otherwise, reversibility can only be of limited use for the complex process of decision-making within climate policy and solar climate engineering, and the development of new technologies more generally.

[36] https://www.whitehouse.gov/ostp/legal/ [last visited October, 2023]

III.2.7 Conclusion

The aim of this paper was to investigate the ethical relevance of reversibility for the evaluation of new and emerging technologies. The case of solar climate engineering through Stratospheric Aerosol Injection served as an explorative example for the relevance and usefulness of reversibility. In the process, irreversibility was linked to the concepts of control, entrenchment, and disruption through NESTs, and how these concepts affect societal structures and institutions. The conclusion of this investigation was that reversibility can serve as a useful tool to better understand the potential for lock-in and disruption through NESTs such as SAI. Nonetheless, reversibility is severely limited in terms of ethical guidance, unless it is clearly contextualized and qualified.

IV. Chapter 3: Systemic Risks and Solar Climate Engineering Research. Integrating Technology Ethics Into The Governance of Systemic Risks[37]

The paper explores how the framework of systemic risks can help govern the risks imposed through solar climate engineering research. The central argument is that a systemic perspective of risk is a useful tool for analysing and assessing the risks imposed through Stratospheric Aerosol Injection (*SAI*) research. SAI is a form of climate engineering that could cool the planet by enhancing its albedo through the injection of aerosols into the stratosphere. Researching such a technology creates systemic risks with a strong sociotechnical component. This component consists of the potential societal harm that a developing or new technology might cause to existing norms, values, institutions, and politics. The systemic risk framework is a valuable heuristic for this case, given the complex interdependencies of societal systems, infrastructures, markets, etc. At the same time, the systemic risk framework can be enhanced through the inclusion of a more robust and reflected ethical considerations on technological risks. Consequently, this article seeks to supplement the systemic risk governance framework with insights from technology ethics. Specifically, the paper offers an ethically reflective conception of societal value dynamism and stakeholder engagement and participation, tying it to existing systemic risk governance approaches.

IV. 3.1 Technology ethics and climate risks

The impacts of climate change are assured to be both harmful and difficult to predict. The Intergovernmental Panel on Climate Change reports highlight this fact, when the potential damage caused by climate change is expressed through a probability scaling of ‘high confidence’ to ‘low confidence’ (IPCC WGI 2023, 4). The future of how climate change will impact the planet is, while clearly harmful for many species, eco-systems, and societies all over the world, also shrouded in risk and uncertainty. Accordingly the assessment and governance of climate risks has become paramount for climate action. One way of governing and abating some of these

[37] This chapter is published as: Hofbauer, Benjamin. 2023b. “Systemic risks and solar climate engineering research. Integrating technology ethics into the governance of systemic risks”, Journal of Risk Research, DOI: 10.1080/13669877.2023.2288010

risks is through new technologies. As with many risks, industrialized societies seek to partially offset or remedy these climate risks through innovation and ever-increasingly complex infrastructures and systems. Climate change, for better or worse is thus often framed as a challenge for technological innovation.

The severity and urgency of climate change has led to expert-driven debates on the research and possible deployment of solar climate engineering through Stratospheric Aerosol Injection (henceforth *SAI*). SAI as an approach entails continuously spraying aerosols e.g. *via* aircrafts, into the stratosphere which would reflect a fraction of the incoming solar radiation. The resulting veil of aerosols would then serve as a temporary sunblock. This method is considered by some scholars as an additional means to slow the rate of planetary warming or even stabilize temperatures, while humanity stops emitting CO2, and eventually achieves progress in the active reduction of atmospheric CO2 (Keith 2013; Crutzen 2006).

However, the process of researching SAI poses a systemic risk to the energy transition through its sociotechnical impacts, i.e. the potential societal harm that the mere possibility of a new technology might cause to existing norms, values, institutions, and politics. From a climate policy perspective, this harm manifests itself as creating societal reactions that run contrary to the overarching goal of climate action, namely the sustainable and just abatement of CO2 emissions, or the risk of 'mitigation deterrence' (McLaren 2016).

To assess and possibly govern this risk, this paper applies existing frameworks of systemic risk governance to SAI research governance. It further expands on these frameworks by supplying additional insights from the ethics and philosophy of technology, as well as risk ethics so far underexplored by the systemic risk community. In contrast to conventional risks, systemic risks are risks that endanger the functioning of vital systems, e.g. infrastructure, supply chains, healthcare systems, or other (Florin and Nursimulu 2018; Schweizer 2021). Framing SAI research as a systemic risk contextualizes the endeavour within the many systems and structures it would affect, most importantly how it could negatively impact the energy transition. What distinguishes this approach from other work on the risks imposed through SAI so far (e.g. Tang and Kemp 2021) is the specific focus on SAI's research process, rather than SAI deployment.

The main argument is that a well-governed research program taking into account the systemic risks of SAI research, could reduce the risk of mitigation deterrence. The systemic risk governance framework can be enhanced in order to account for the dynamic nature of societal values and the moral relevance of participation when evaluating New and Emerging Sciences and Technologies (*NESTs*). This requires a theoretically grounded conceptualization of both participation and societal values in the context of NESTs, which this paper provides. Ethics and philosophy of technology scholarship can supplement existing systemic risk governance frameworks, providing reflective insights about the interdependent relationship between technology research and societal values, and establishing a more rich understanding of participation as a means to achieve justice in a democratic decision-making process.

The paper first connects societal risks to technology, exploring how the two are reciprocally related in modern industrialized societies. On this basis, the risks invoked through SAI research are presented in the subsequent section. By way of applying the systemic risk framework to the case of SAI research, two supplements for this framework are proposed in the form of societal value dynamism and stakeholder engagement and participation. These two supplements are argued to create a more politically acceptable and ethically justifiable form of governing SAI research's risks. The concluding section reiterates the central findings and suggests space for future investigation with regards to risk governance for New and Emerging Sciences and Technologies in general.

IV. 3.2 Contextualizing societal risks and technology

Modern, industrialized societies often rely on technological innovation to handle the unwanted side effects of daily business, facilitate certain activities, or explore new possibilities for action. This inevitably raises ethical questions. Whether and to what extent we should rely on technology and innovation, and what kinds of technologies we ought to research and build, are all normative questions that require moral reasoning and political engagement. New technologies often lead to new possibilities for action, resolve existing problems at the cost of causing new ones, and add to the complexity of existing infrastructure and sociotechnical systems. Electric vehicles for example resolve the problem of pollution and direct CO2 emissions of traditional combustion engines, while also causing problems for the supply of electricity, and the just and environmentally friendly extraction of rare metals. Additionally, they require a modification of the existing mobility infrastructure in the form of new charging stations, as well as large scale transmission cables.

Due to the complexity and vastness of current societies, made up of interrelated markets, politics, institutions, infrastructures, etc. it would be imprudent for researchers, policy-makers, and the public to try and anticipate all the outcomes of researching and introducing New and Emerging Science and Technologies (*NESTs*). Rather, exploring and evaluating the impact of potential NESTs has been framed as a dynamic process that puts specific focus on possible harms and benefits, i.e. risks and uncertainty. From an ethical perspective, this has led to increasing attention being paid to the assessment of risk imposition related to technological development in the past years (Hansson 2017; Taebi 2017; Grunwald 2015), with various focal points on e.g. public acceptance, policy-making, participation, and others. The conceptual aim of risk ethical assessment is to understand the moral implications of imposing risks (e.g. as domination, see Maheshwari and Nyholm 2022), and the possible justification for the imposition of risks through technology (Hansson 2013).

It is through this lens of continuous risk imposition, assessment, and abatement that I analyse SAI research's risk of leading to harmful consequences such as 'mitigation deterrence', i.e. to a reduction in societal willingness to mitigate. Consequently, it is first important to take a closer look at the interrelated nature of societal values and technological risks. I refer to the continuous innovation and deployment of new technologies constantly reshuffling the imposition of individual, communal and societal risks as the *sociotechnical risk cycle*. The sociotechnical risk cycle is a schematic process and is neither an exhaustive nor exclusive means by which such risks can be explored. The following example describes this cycle.

The flammability of a building might be drastically reduced through the introduction of some building material. However, at a later point in time research and empirical data indicate that this building material is carcinogenic and thus increases the likelihood of workers developing lethal diseases. Consequently, a new material is developed that is both less flammable and not carcinogenic. Unfortunately, this new material turns out to be extremely damaging to the environment due to the materials necessary for its production. From a schematic perspective, a technology was introduced, a given risk was identified on the basis of empirical insights and value claims (e.g. health and the environment ought not be harmed), and subsequently abated through the introduction of a new technology. Finally, the new technology created novel or different risks, and the constant process of risk identification – abatement – and novel identification repeats itself.

Ulrich Beck describes a similar process in his seminal work titled the *Risk Society* (1992). Beck specifically highlights the relationship between producing and increasing wealth and the risks that this production incurs (Beck 1992, 19):

> In advanced modernity the social production of *wealth* is systematically accompanied by the social production of *risks.* Accordingly, the problems and conflicts relating to distribution in a society of scarcity overlap with the problems that arise from the production, definition and distribution of techno-scientifically produced risks.[1]

Note that this cycle is difficult to break, not least because a continuous growth paradigm seems to undergird any reason for political and economic action in industrialized capitalist societies. With regards to a new technology, even if it resolves past problems it will introduce some form of novel or known risks. This is due to the inherently dynamic characteristic of risk in a value pluralist and dynamic society that is open to the emergence of new values and norms, as well as the overhauling of existing norms and values. Values, in this sense refer to 'what a person or group of people consider important in life' (Friedman, Kahn, and Borning 2008, 70). They are ideals that people seek to realize.

The sociotechnical risk cycle and its persistence hinge, I argue, on the following three central characteristics of risk. One, the identification of risks is both the result of empirical observation and moral evaluation. In the case of building materials, we find that different values were at play at different moments in time. First, the safety and security of houses, second the health and well-being of the workers, and third some form of ecological integrity. In short, the assessment of those risks was based on specific values at each moment in time, which shows how risks are value-laden phenomena. The value-laden aspect of risk is also acknowledged in the sociological tradition of risk assessment. Thus when risk scholar Ortwin Renn describes risks as being something that has 'an impact upon what humans value' (Renn 2008, 2) some form of moral argument needs to define what that human value is. What constitutes a value is an ethical question that cannot be answered through observation alone.

The fact that identifying risks is a political and ethical process highlights the need for meaningful participation that accounts for historical injustices and structural inequities. As the working material example demonstrates, there needs to be political action claiming that the environment is worthy of protection, and therefore harming it constitutes a risk. For this political action to be effective however, a decision-making process is required that allows for meaningful participation among all stakeholders, and potentially remedies existing power imbalances. Risks, defined as potential harms accordingly cannot but rely on an evaluative premise and the identification of risks is always also a political process.

A second characteristic that the above example highlights is the constant reshuffling of risks in liberal, industrialized societies. The possible harm to one value (e.g. safety) is abated through the introduction of a new material. However, the intervention does not end here, since new values emerge, or existing values become more prominent. A labour union might be politically active and demand better conditions for their workers, which entails not having them be subject to carcinogenic substances. Thus the value of individual and communal health and well-being, perhaps also the value of social equity in terms of how workers are being treated, becomes a new prominent factor for the existing risk assessment. This dynamic development of values is inevitable in value pluralist societies, wherein the public opinion on what should and should not matter can become political and institutionalized, and thus leads to and requires a continuous reassessment of existing risk-impositions. This dynamism is a central challenge for the implementation and maintenance of technology and infrastructure, as these structures themselves are the manifestation of certain values that may change over time. For example, twentieth century flood risk management systems that exclusively focus on safety and ignore their environmental impacts, may no longer be deemed adequate by the public due to changing values (Taebi, Kwakkel, and Kermisch 2020).

The third and final aspect is the way in which risks are often abated or mitigated in industrialized societies, namely through technological innovation and intervention. This third aspect is helps understand the relationship between societal values, technological development, and risk. Actions and behaviours on a societal level are often not fundamentally changed just because society realizes that these behaviours carry risks. Rather, since these behaviours are manifestations of specific values or notions of a good life, applied science and policy seeks ways to mitigate or abate the risk through technological innovation. People may not want to stop insulating their housing units, since it is assumed that insulated housing is a valuable action (to save money, to protect the environment, etc.) Instead of adapting or changing our behaviour (and the values that drive them), we seek to sidestep the risks that behaviour induces by changing their context. This is what some scholars have framed as the so-called 'techno-fix' (Huesemann and Huesemann 2011; Sand, Hofbauer, and Alleblas 2023; Jongsma and Sand 2017), i.e. seeking to resolve societal ills through technological innovation.

The persistence of the sociotechnical risk cycle is particularly glaring in the context of climate change. This persistence invites the claim that humanity cannot innovate its way out of climate change, since the constant reliance on innovation based on an assumption of continuous growth, are drivers rather than solutions for the ongoing ecological crisis. Innovation often ignores the fundamental structural causes for climate harms, while doing little to alleviate climate injustices. The way climate change harms vulnerable communities in particular, is often tied to the precarious context within which these groups and individuals find themselves, which in turn are co-products of underlying marginalization through political oppression, exploitation, and/or discrimination. The solution to these injustices is not technological development, but societal change found in the politics of e.g. degrowth or deep ecology movements (Kerschner et al. 2018).

At the same time, categorically dismissing NESTs as a potential additional means to combat climate change comes with its own ethical and practical challenges. From an ethical perspective, it is not a given that NESTs necessarily lead to more climate injustice and further marginalization. Rather, it depends on how these technologies are envisioned, and what role they are imagined to play in the energy transition. Decentralization, a main feature of the degrowth paradigm, might be advanced through innovation (Pesch 2018). Further, overcoming climate change and the ecological disaster accompanying it, necessitates swift, sustainable, and effective intervention. As the recent IPCC report shows, there is no chance of achieving the 1.5 or 2 degrees Celsius goal without the implementation of Negative Emissions Technologies (IPCC WGI 2023, 184):

> Past, present and future emissions of CO_2 therefore commit the world to substantial multi-century climate change, and many aspects of climate change would persist for centuries even if emissions of CO_2 were stopped immediately (IPCC 2013b). According to AR5 [the fifth Assessment Report], a large fraction of this change is essentially irreversible on a multi-century to millennial time scale, barring large net removal ('negative emissions') of CO_2 from the atmosphere over a sustained period through as yet unavailable technological means.

Accordingly, the climate crisis that humanity has produced now requires all hands on deck to resolve, or at the very least make it the least catastrophic we can. The Paris goals are also closely linked to reaching certain, irreversible climate tipping points that would have catastrophic impacts on the biosphere as a whole (Armstrong McKay et al. 2022). In addition, current main-stream analysis of climate change, such as the IPCC reports, tend to underrepresent the more dangerous and potentially even more catastrophic climate scenarios (Kemp et al. 2022). The realization of climate change's irreversible impacts and that it could turn out much worse than anticipated underscores the need to consider an array of interventions, including, but not limited to technological approaches.

IV. 3.3 Stratospheric Aerosol Injection research and sociotechnical implications

Given the urgency for action that climate change causes, some scientists and experts have explored the possibility of directly influencing the climate through technological means. One form of doing this is through Stratospheric Aerosol Injection (*SAI*), which entails continuously spraying aerosols e.g. *via* aircrafts, into the stratosphere. These aerosols would create a veil that could temporarily block some of the incoming sunlight. The central premise of this intervention is that it could serve as an additional means to slow the rate of planetary warming, giving humanity more time to eventually reduce atmospheric CO_2 (Keith 2013; Crutzen 2006).

The potential of a fully-fledged, global research program on SAI has led to heated debates throughout the natural and social sciences dealing with climate change. Some scholars from a variety of fields have publicly called for a moratorium on deployment (Biermann et al. 2022), which prompted others to push for a 'balanced' account of SAI and its research process (Wieners et al. 2023). Calls against SAI research usually point out that researching a technology inevitably leads to some form of deployment. This is commonly referred to as a 'lock-in' or 'slippery slope' occurring, wherein political or economic incentives outweigh the scientific and justice related rationales when deciding whether to go ahead with the development of a certain technology.

Applied to SAI, the research runs the risk of inevitably leading to deployment despite the scientific community's best intentions (McKinnon 2019; cf. Callies 2019).

Another important warning against SAI research is that investing resources and political capital into such a risky and potentially unjust form of climate mitigation is both imprudent and morally problematic (Biermann et al. 2022; Hamilton 2013, 2014; Schneider 2019). In contrast, those defending the idea of a globally implemented research project point towards the catastrophic climate situation humanity already finds itself in and argue that we cannot make a well-informed choice about SAI if we do not understand it well enough (Wieners et al. 2023; Winsberg 2021). Accordingly, only a well-coordinated SAI research process with proper political and ethical governance can tell us whether the technology is viable and justifiable.

Following this latter line of reasoning the argument developed here assumes that a research ban is in and of itself problematic, since it presumes that technology development is the direct application of scientific knowledge and forecloses any risk or uncertainty. It relies on a potentially deterministic understanding of technological development, which seems to unduly simplify the way new technologies are developed, and their sociopolitical embeddedness.
With this assumption in mind, I focus on the research process of SAI, and make no claim about deployment. The evaluation of SAI research lends itself to an ethical risk assessment, since its outcomes, worries, promises, etc. are all shrouded in a considerable amount of uncertainty. In other words, most of what we can say about SAI research comes in the form of possible outcomes, possible harms, and possible benefits. However, traditional risk analysis and governance approaches seem ill-equipped to handle the inherent complexity of the technology-climate change-society nexus. I believe that a systemic perspective of risk, which focuses on the complex, interdependent nature of societal and technological risks, is a useful tool for analysing and assessing the risks imposed through SAI research better and make space for their ethical evaluation.

IV. 3.4 Systemic risks & technology: the normative relevance of values

The concept of systemic risk has its origins in economics and finance. Scholars such as Kaufmann and Scott for example explore the concept when discussing the potential contributors to the breakdown of entire financial systems, and the subsequent systemic effect of that breakdown (Kaufman and Scott 2003). With reference to the OECD report on systemic risks (2003), the International Risk Governance Council's released a report specifically focusing on their governance, further characterizing systemic risks in lieu with its 'cascading effects', highlighting the interdependent and often non-linear relationship of triggers and outcomes (Florin and Nursimulu 2018, 9). Risk scholars such as Ortwin Renn and Pia-Johanna Schweizer have substantially reworked and deepened the concept, underscoring its relevance for the challenges humanity faces in the twenty first century (Renn et al. 2022; Schweizer and Renn 2019; Schweizer 2021). It is specifically this latter approach, with its focus on transdisciplinary and qualitative assessment of systemic risks that I focus on.

Systemic risks are risks that endanger the functioning of any vital system, e.g. infrastructure, supply chains, healthcare systems, and other systems central to the functioning of society. They can be contrasted with traditional risks through the following characteristics.[2] Their cause-and-effect structure is non-linear, which means that apparently minor impacts may have severe outcomes for the system. This non-linearity also leads to system lags and tipping points, wherein the consequences of a given input may only manifest themselves at a later moment in time. Further given the complexity of the systems affected through systemic risks, there is a considerable amount of uncertainty in their assessment, and consequently their governance. This complexity exacerbates the high and latent interdependency of both the system under investigation as well as the mechanisms that potentially led to a systemic breakdown.
While Renn and colleagues mention five aspects of risk governance in total, I specifically focus on two of them, namely (1) the need for 'recurrent, adaptive, and synchronized' governance (Renn et al. 2022, 1914) and (2) the epistemic importance of meaningful participation (Renn et al. 2022, 1915). I believe that these two aspects can be most illuminatingly enhanced through the inclusion of ethical considerations. They subsequently give an indication of how systemic risks governance can be further ethically grounded, ensuring a reflective and adaptive assessment process, and having a moral basis for the broad participation of all actors and agents involved (see Renn et al. 2022, 1913ff.) Importantly, enhancing these two aspects by actively incorporating ethical reflection into the participation and decision-making process presents a valuable heuristic when setting up a holistic SAI research framework.

The first characteristic that Renn and colleagues discuss is the necessity for systemic risk governance to be ‘recurrent, adaptive, and synchronized’ (2022, 1914). Iterative and adaptive approaches are central to any form of decision-making under risk and uncertainty (Marchau et al. 2019; Taebi, Kwakkel, and Kermisch 2020). Governance frameworks that have been proposed for SAI research so far all include reflexivity, adaptability, and iteration as central tenets of the research process (NASEM 2021; Gardiner and Fragnière 2018; Bellamy 2016; Hofbauer 2023). The need for dynamic governance can be most easily explained through the fact that the identification of risks goes in tandem with the gathering of new information. Empirical findings put the governance process into a new context, e.g. when scientific inquiry leads to the realisation that SAI may lead to a change in precipitation patterns in highly rain dependent regions (Niemeier and Timmreck 2015; Tilmes et al. 2020).

However, dynamic governance is also a highly relevant framework for the ethical assessment of risks. As the sociotechnical risk cycle highlights, introducing new technologies entails the continuous abatement of existing risks and the related production of novel risks. Since risks are also the product of value judgements, a novel risk might not just be the outcome of new empirical findings, but actually related to societal developments that re-evaluate an existing state-of-affairs on the basis of new values. These new values might themselves be a reciprocal product of, or influenced by the new technology. Importantly, new values might arise and impact the research of new technologies before these technologies have been fully deployed or developed. Hence new values create and are being created by the interaction of societal norms and values, as well as the infrastructure and technological ‘imaginary’ that a given society finds itself in Jasanoff and Kim (2015; Hilgartner 2015). This is mainly what scholars sceptical of SAI research fear, when they claim that such a program would inevitably lead to more exploitation and injustice. They argue that researching SAI is inherently morally wrong, since its goal is hubristic and its very concept incompatible with values such as democracy and global justice (Hulme 2014; Szerszynski et al. 2013; Schneider 2019). This potentially harmful impact of SAI research, in turn, poses as a systemic risk towards a just energy transition, undermining justice considerations in favour of a quick technofix.

An SAI research program might thus influence specific values that a society holds, such as how sustainability and justice relate to one another (Hofbauer 2023). A research program that accounts for its potential impacts on existing sociotechnical imaginaries consequently needs to be aware of its framing and narrative, and proactively lead the charge in how its technology is communicated. When scholars such as Jebari et al. (2021) argue that SAI research proposals need to come in tandem with thorough mitigation measures, they also address the need for proper framing: SAI research, whatever its outcome will not resolve climate change. This is of course an immense challenge, as narratives surrounding technology are not solely in the hands of scientists, but rather a complex compound product of political interests, institutions and societal expectations (Borup et al. 2006). Nonetheless an integrated mechanism for the framing of SAI could help provide a common context for the conversation surrounding its research.

So far, systemic risk scholarship has paid little attention to the reciprocal nature of such risks arising in the society-technology nexus, i.e. societal risks that emerge through the introduction or research of NESTs. Importantly, these risks should be considered an additional form of systemic risks that require ethical analysis and reflection. Assessing these risks entails not just accounting for the empirical observation of societal values, but also for the normative fluidity of existing and future risks and their interdependence with NESTs. Actively incorporating framing mechanisms that highlight the limitations of SAI research, for example, could serve as a tool to account for the risk of setting problematic expectations or reinforcing existing unjust worldviews.

IV. 3.5 Risks & technology: a procedural and substantive account of participation

The second characteristic of systemic risk governance that can be supplemented with ethical considerations is inclusivity and the broad and meaningful engagement of stakeholders. As Renn and colleagues explain, inclusivity is paramount from an epistemic point of view. A more diverse set of perspectives will help better identify the risks at hand. Referencing Florin 2013, Renn et al. argue that (2022, 1915):

> Having many stakeholders involved provides a much more effective guarantee to pay attention to a multitude of early warning signals and to detect irregularities that may be outside of the screen that official risk observers use.

In other words, there is epistemic value in including all kinds of actors involved and affected by the system, especially in terms of making sure that all risks are accounted for, and that the system is properly understood. With the term epistemic I refer to the idea that stakeholder participation can teach researchers and policymakers about the values stakeholders hold. Again, this example shows how values come into play and how e.g. the concept of participation might play a role for the identification of relevant risks.

In an earlier piece, Renn and Schweizer discuss in-depth not just the epistemic relevance but also the normative importance of inclusive governance approaches for policymaking in complex, systemic risk situations (Renn and Schweizer 2020). Focusing on the societal implications of the energy transition, the authors write that '[…] procedural structures are urgently needed that build upon the best available expertise and the informed consent of those who will experience the consequences of the requested changes'. (Renn and Schweizer 2020, 41) Accordingly, the authors refer to some ethical considerations regarding participation, such as Habermas' ethics and discourse theory (Habermas 1984), to highlight the need for participation as an ethical criterion.

To clarify the ethical relevance of participation, it is useful to distinguish between procedural and substantive approaches to ethics and questions of justice.[3] In a nutshell, while substantive accounts of justice focus on the outcome of a given decision-process, i.e. whether the final agreement or distribution is just, procedural accounts of justice emphasize the importance of inherently just processes.[4] The inclusive governance of systemic risks relies on a robust conception of justice as participation. This means that justice is conceptualized within the boundaries of fair procedures alongside a set of substantive ethical guardrails. This is what the political philosopher John Rawls referred to as 'imperfect procedural justice' (Rawls 1999 , 74f.). Imperfect procedural justice describes a decision-process wherein the outcome of said process is not immediately just merely by virtue of having followed a specific procedure. It also requires evaluating the decision's outcome. In other words, broad participation alone does not ensure a just outcome.

The focus on procedural justice is motivated by two central characteristics of decision-making in the systemic risk context. First, systemic risks are far-reaching, interdependent phenomena that affect a variety of stakeholders with diverse, often incompatible sets of values or interests. Meaningful and reflective participation provides a stage for the assessment, deliberation, and consolidation of those values, as well as a ground to establish guardrails against potentially illegitimate value-claims.

Second, systemic risks are often shrouded in uncertainty and ambiguity, reducing the predictive capacity and epistemic authority of expert-driven deliberation, i.e. inductive reasoning through the scientific method. The iterative nature of learning about the systemic risk phenomenon requires a dynamic and reflexive process of assessment, evaluation, and deliberation to account for both the discovery of new information and how that new information affects existing value-judgements by creating a new context within which these judgements are made.

Accordingly, the risk of SAI research resulting in a technocratic technofix, leading to mitigation deterrence with no underlying justice or sustainability considerations can be clarified and accounted for through this procedural and substantive distinction of justice. Robust and meaningful participation provide both useful insights and critical oversight about the research processes, ensuring that value considerations relevant to e.g. the most vulnerable stakeholders are an integral part of the decision-process. Further, clear guardrails in terms of which stakeholders get to have how much say, and what degree of potential environmental damages are acceptable provide a kind of deliberative playing field for the procedure, while also accounting for critical substantive values.

IV. 3.6 Contextualizing participation

Building on the ethical considerations for inclusive governance, participation and recognition can best be understood as a central tenet to achieving justice (Fraser and Honneth 2003). Importantly, meaningful participation is a catalyst to frame all stakeholders as meriting equal recognition of the potential harms a NEST might cause for them, while also being able to historically contextualize and differentiate between different kinds of needs and harms. This combination of equity and differentiation is particularly central to a NEST of global impact such as SAI (Hourdequin 2019) and plays a major role in the ethical evaluation of a just energy transition (van Uffelen et al. 2022). Similarly, the labour union that looks out for the well-being and health of its members is a crucial stakeholder in the sociotechnical risk cycle explored above, and their participation in the decision-making process is ethically indispensable. Stakeholder engagement then is not merely prudent and epistemically valuable, but necessary if the risk assessment and governance process is to be morally justifiable.

Unchecked and decontextualized participation might lead to unjust processes as well. When trying to deal with climate change and its accompanying harms, uncritically giving the same weight to all the actors involved runs the risk of distorting procedural and recognitional aspects of the decision-making and governance process. While it is not always possible to pre-emptively assess the intention of all stakeholders, putting their role in the system at risk into context and taking their past actions into account can give a more holistic and reflected understanding of their intentions. Further, some actors may have incentives to stall rather than enable radical climate action, which means that the role and motivation of these actors needs to be contextualized, and the stakes they hold evaluated. For example, the stakes of keeping a fossil fuel economy at a profitable level are potentially morally less relevant (perhaps even immoral), as say the stakes of future generations to inherit a habitable planet.

This difference in stakes further influences the kinds of risks we might deem acceptable, and their reasons (e.g. in the form of values) ought to be central and transparent to the debate on risk governance. Transparency serves not only the epistemic purpose of knowing who holds which values, but also a moral and political purpose for when society is confronted with the question of how the different risks from different stakeholders are to be evaluated and taken into account. The point of this broader understanding of participation then is not to pre-emptively judge the stakeholders' position, but to give all stakeholders a reasonably informed forum for a meaningful debate on the values and risks involved in the decision-making process surrounding a NEST. An incorporated ethical reflection helps contextualize the arguments from different stakeholders based on their underlying values, which in turn could serve to set guardrails against potentially unacceptable risk impositions.

In this sense participation is a sine qua non condition for justice, while at the same time requiring some guardrails and limitations on the kinds of risks that should be considered and form part of the decision-process. Participation in the form of actively reflecting on the values involved in the arguments raised by stakeholders further highlights the importance of understanding the dynamic relationship of the risks, values, and technology. A dynamic account of sociotechnical value change underscores the fact that continuous participation is necessary, since values may change over time and through the NEST introduced. Actively including and accounting for the dynamic and interrelated process of changing societal values and technological development, as well as the ethical relevance and limitations of participation broadens the existing systemic risk governance framework. Consequently, the framework becomes more holistic when evaluating NESTs, or even their research process such as SAI research governance proposals. Such a holistic and systemic framework would further allow us to take a nuanced perspective on the ethical underpinnings of governing SAI research processes, and the moral and sociopolitical challenges such a research process needs to take into account.

IV. 3.7 Conclusion

Any SAI research program needs to inherently and explicitly address the question of how it aims to deal with the risks it produces, such as mitigation deterrence effects, the risk of unilateral decision-making, and the unreflected inclusion and exclusion of different values. The non-reflective, and decontextualized introduction of an SAI research program is a systemic risk for the energy transition, potentially inhibiting the necessary shift away from fossil fuels. SAI research carries the real risk of undermining the abatement of emissions through indirect technological means, based on expectations and expediency, rather than scientific facts. However, this risk can possibly be countered through the introduction of a number of measures, on both the side of SAI research, and the impacted system, i.e. the energy infrastructure. The systemic risk framework is insightful to account for the complex nature of risks imposed through SAI research, specifically providing a systemic outline of the interdependent structures and systems at play. At the same time, the framework can be more ethically grounded through additional reflections on the value-technology relation as well as the justice requirement of participation. This could be done *via* e.g. a more holistic policy framework that already includes measures beyond merely researching SAI, but also addressing the risk of deterrence effects (Jebari et al. 2021). Further, such a framework could help tie any SAI research effort explicitly to binding mitigation commitments and decarbonization policies. On an international level, this would entail carving out clear responsibilities for nations, but also corporations and industry. Further, policy measures such as risk-response feedback frameworks (Jebari et al. 2021), anticipatory governance approaches, as well as assessment methodologies that account for the risks and uncertainties involved in such a research process can anticipate ways of dealing with the potential of mitigation deterrence. On this basis, it is possible to ethically govern the systemic risks associated with SAI research.

Notes

1. Beck points out to what degree industrialized societies create new inequalities through unequal of risk impositions (1992, 100).
2. This section largely follows and builds on Renn et al. (2022), 1904 ff.
3. I would like to thank an anonymous reviewer for inspiring me to add this fruitful distinction.
4. There remains much more to be said about this distinction. For an insightful overview, see Miller (2023).

V. Chapter 4: Techno-Moral Change Through Solar Geoengineering[38]

This article brings a new perspective to the ethical debate on geoengineering through stratospheric aerosol injection (SAI), incorporating the emerging techno-moral change scholarship into the discussion surrounding sustainability. The techno-moral change approach can help us understand different ways in which technology might shape society. First, it helps highlight how values and norms are interrelated. Second, it shows that techno-moral change can happen even if the technology is in no way realized. Through the introduction of two techno-moral vignettes, two diametrically opposed ways in which SAI forces us to rethink sustainability and our relationship with nature are suggested. SAI could lead to a situation of entrenchment, wherein sustainability as a norm is undermined, or transformation where the necessity of acting according to sustainability is highlighted.

V.4.1 Introduction

Rapid global warming and climatic change have impactful consequences for nature and humanity. Slow political action and insufficient policies, combined with the increasing and unpredictable threat of climate change challenge current ways of living. Industrialized societies appear to have ended up in an unsustainable situation (i.e., the resources consumed by human beings are out-weighed by their rate of reproduction).

To overcome this unsustainability, scientists, scholars and policymakers are consider-ing technological solutions in the wake of realizing just how far-reaching human impact on the climate and the environment is. The idea is that technology can help societies be more sustain-able, as the European Green Deal exemplifies with its heavy focus on research and development of green energy production and storage (European Commission, 2019). Framing climate change as a technological problem rather than a behavioural one evokes the term 'techno-fix' (Weinberg, 1966). Since it is difficult to implement behavioural changes through policy, soci-

[38] This chapter has been published as: Hofbauer, Benjamin. "Techno-Moral Change through Solar Geoengineering: How Geoengineering Challenges Sustainability." *Prometheus*, June 1, 2022. https://doi.org/10.13169/prometheus.38.1.0082.

etal issues are supposedly ‘fixed’ through technological development (cf. Huesemann and Huesemann, 2011). Achieving sustainability becomes a purely technological task, rather than a societal one.

One particularly prevalent example of this is the idea of geoengineering, intentionally manipulating the earth’s climate through technical means (Keith, 2001). There are different ways of engineering the climate (for an overview, see Shepherd *et al.*. 2009). However, focus here is on one particular form, namely stratospheric aerosol injection (SAI) (Crutzen, 2006; Keith, 2013; Tilmes *et al.*, 2020). In short, SAI entails spraying aerosols into the stratosphere, which reduces some of the incoming sunlight, cooling the earth’s surface temperature. Some consider the technology as a possible way of buying time to reduce parts of the most dangerous climate change impacts in the upcoming decades. At the same time, it is generally agreed that the technology is not a silver bullet solution to climate change (Keith, 2013; Horton and Keith, 2016; Svoboda *et al.*, 2018).

The topic of SAI has received growing attention from the fields of science, technology and society studies, ethics and political philosophy in recent years (for an overview, see Stilgoe, 2015; Pamplany *et al.*, 2020). Throughout the debate on the ethics of SAI, a specific focus is usually on issues of historical, procedural and distributive justice (Preston, 2016). If sustainability is taken into account when discussing SAI, it is in order to make a point about its unsustainability (Adloff and Hilbrich, 2021), or how the sustainability debate can bring different viewpoints on SAI together (Thiele, 2019). How SAI potentially affects conceptions of sustainability, however, is underdeveloped. This article seeks to fill the gap on how SAI as technology can lead to techno-moral change (Swierstra *et al.*, 2009). Focus is on the norm of sustainability and how SAI affects this norm and the underlying value of nature.

V.4.2 Techno-moral change

Technology and its development have an impact on society, morality and the realm of the political, as discussed throughout the philosophy and ethics of technology (Winner, 1980; Verbeek, 2011; Grunwald, 2015). Over the past decade, scholarly interest has focused on how technologies mediate and influence ethical norms, values and concepts. Such a focus entails an emphasis on co-dependent evolution, where society evolves with the technologies it creates. This field of research is commonly referred to now as ‘techno-moral change’ (moral change brought about through technological development) (Swierstra *et al*., 2009; Swierstra, 2013; Nickel *et al.*, 2022).

The scholarship of techno-moral change (TMC) is based on Deweyian pragmatism and the field of science, technology and society (STS) studies, with the aim of including normative considerations into them (Keulartz *et al*., 2004). In pragmatic fashion, TMC opposes ethical foundationalism. It does away with the idea that an ethics of technology should seek universal moral values or truths. Rather, TMC highlights the dynamic character of modern, pluralist societies, which are constituted by continuous technological development, and zooms in on how technology impacts moral values and norms (Boenink *et al.*, 2010; Swierstra and te Molder, 2012).

To understand better what precisely changes when TMC occurs, Tsjalling Swierstra's framing of morality as a force field is helpful (Swierstra, 2013). In a pluralist society, there is a constant underlying debate on what kinds of norms and values, which are different subsets of morality, should be acceptable. As an example, we can think of how Roman Catholic morality competes with secular feminist morality, both of them with their respective norms and values. TMC occurs when the introduction of a new technology tips the scale of the competition in favour of one morality. The norm of chastity and the value of virginity lost footing in many societies, in part, as a result of the technological development of the contraceptive pill.[39]

At heart, techno-moral change occurs through the coevolution of technology and society (Rip and Kemp, 1998). Technology influences societal norms, for instance, by making 'some moral options easier to argue for and others less easy', as the example of the contraceptive pill shows (Swierstra *et al.*, 2009, p.120). By opening up new courses of action and closing off others, technology changes how one can behave in a given situation. Technology might make it easier to do the right thing by changing our range of options or motivations. This is most explicit in so-called 'behavioural change' technologies that give us direct feedback on our actions, and try to nudge us into certain habits (Frank, 2020).

[39] Of course, voices of change, resistance and social movements are at least as, if not more, important for moral change (Anderson, 2014; Baker, 2019).

Similarly, technologies can afford and constrain actions that realize certain values (Brey, 2005, 2018). For example, prenatal diagnostics allow parents to know about the state of health of their child before giving birth, affording them a range of previously unattainable moral options. Thus, whether to give birth to a baby with certain health conditions becomes a choice rather than the acceptance of fate. Accordingly, technology mediates how we perceive the world and ourselves and alters the range of possibilities for action and deliberation. Ultrasound imaging positions expecting parents as decision-makers who deliberate on the basis of the expected health of their unborn child (Verbeek, 2011). The technology forces a whole range of new ethical questions onto the parents (e.g., on what basis to evaluate whether a child is worth having, whether to know, or want to know, about potential health concerns).

Realizing that technology actively shapes morality is insightful when it comes to the evaluation and analysis of existing technologies after their deployment. This is especially interesting since norms and values are not isolated from one another, meaning that a change in one of them often entails a change in others (Swierstra, 2013). This interdependence of norms and values is best showcased by unforeseen consequences of new technologies. Again, one can look into how the introduction of the contraceptive pill supported the feminist movements in the 1960s and 1970s (Baker, 2019; Nickel *et al.*, 2022). While the pill allowed for more sexual freedom, it might simultaneously be argued to have shifted the burden of taking care wholly onto women in terms of unwanted pregnancies. Whether and how contraception is made technically possible accordingly entails new possibilities – and burdens – for a society and its individuals.

Another interesting example of techno-moral change is the introduction of mechanical ventilators during the 1950s. This technological innovation caused widespread moral uncertainty about whether withdrawing ventilation from brain-dead patients is morally permissible since it was unclear whether this would constitute murder (Baker, 2013, 2019; Nickel, 2020). Before the use of mechanical ventilation, a person was simply considered to have died when the brain no longer supported breathing. As a result, the heart would stop beating, leading to death. However, it was possible to continue circulation and a beating heart through external means with this new technology. Thus, while the patient's brain stopped functioning, the heart was still beating, leading to the necessary introduction of a new concept of between life and death – brain death. This example highlights how the introduction of new technologies can be disruptive to a set of moral norms, calling into question concepts previously thought to be relatively stable, such as death.

However, it is arguably insufficient from an ethical point of view to understand the moral implications of a technology only after its implementation. Especially in terms of policy-making and governance for new and emerging technologies, it is even more critical to understand the potential impact *ex ante*. The burning question is to understand how research and deployment of certain new technologies affect morality and underlying societal norms before deploying them. In this vein, Taebi and colleagues have explored a framework to account for the development and deployment of new and emerging technologies in the face of the climate crisis (Taebi *et al.*, 2020).[40] Specifically, they focus on how different kinds of uncertainties can arise, such as how public values might change over time.

Accounting for uncertainty thus acts as a safeguard in dealing with all possible kinds of futures. Developing taxonomies and frameworks that consciously include the uncertainties surrounding new and emerging technologies is only one aspect of the solution. The other aspect of dealing with techno-moral change is through the building of scenarios, imagining how the future might change in terms of values and norms. Anticipating or predicting these and how the future will play out is a near impossible task. This holds true also for the effects of technological innovation on society, in which case the relationship between prediction and deployment is in itself a vexing issue. The time when a technology is most malleable (i.e., design and development) is also when its impacts are most uncertain. Once the impacts manifest themselves through the gradual introduction of the technology, its malleability is drastically reduced since it is difficult (if not impossible) to change retrospectively anything about the technology (Collingridge, 1980). This interrelation between uncertain futures and technological development can be found throughout the ethics of technology and geoengineering in the form of the Collingridge Dilemma, lock-in or path-dependency (Cairns, 2014; Genus and Stirling, 2018).

The apparent predicament of not being able to predict the future, however, does not entail openness towards all possible futures. How the future develops is itself based on values and ideals of what should and should not be part of the future. What events or dangers call for pre-

[40] I am indebted to an anonymous reviewer for pointing me towards this paper.

emption or precaution, or what society needs to prepare for, are all present considerations that shape the future (Anderson, 2010).[41] Societies are based on imaginaries that project ways of living from the present into the future (Taylor, 2004; Jasanoff and Kim, 2015), and expectations of how the future should or will turn out manifest themselves in the present through experts, research programmes and institutions (Hilgartner, 2015; Sand, 2018). The future is not a value-neutral field of possibilities, but a space of negotiation for possible scenarios of moral change.

With this in mind, predicting a technology's moral impact becomes less important for its evaluation. Instead, one can focus on how the competing values and norms might be influenced by technological development through a process of scenario building, or techno-moral learning (Swierstra, 2013). In order to understand the impacts that nanotechnology might have at a societal level, Swierstra (2013, p.216) writes that we need to 'train our capacities for techno-moral imagination by developing scenarios and vignettes'. This form of scenario building can help anticipate the societal impact of new and emerging technologies, adding potentially fruitful new points of view to a decision-making process otherwise difficult to determine. The goal of such imagining is not to predict outcomes but rather to illuminate possible future scenarios. Although these moral scenarios might not occur as imagined, certain aspects of them can nonetheless help us understand what kind of values and norms a new technology influences and how.

V.4.3 Solar geoengineering through stratospheric aerosol injection

Geoengineering can be understood as an attempt to deliberately counteract the human impact on the climate through technological means (Keith, 2001). There are several ways of engaging in geoengineering, ranging from high technology proposals, such as mirrors in space, all the way to seemingly natural or harmless solutions. such as increased afforestation or painting rooftops and roads white (Shepherd *et al.*, 2009).

Since the Royal Society's report on geoengineering, it has become commonplace within the literature to distinguish between two main strands of geoengineering, namely carbon dioxide reduction (CDR) and solar radiation management (SRM) (Shepherd *et al.*, 2009). The main difference between the two approaches is that CDR aims to reduce the current amount of

[41] I would like to thank an anonymous reviewer for making these connections.

atmospheric carbon, hence the removal. By contrast, SRM approaches seek to increase the planet's reflection levels, reflecting more sunlight back into space before it can heat up the atmosphere.[42]

One way of achieving SRM is by spraying aerosols into the atmosphere – so-called stratospheric aerosol injection (SAI) – which would increase the earth's reflection levels and subsequently reduce surface temperatures (Niemeier and Tilmes, 2017; Shepherd *et al.,* 2009). SAI, therefore, represents a technology that could potentially aid humanity in the fight against climate change.

With this backdrop in mind, the question of how a technology such as SAI can impact current societal norms and values arises. SAI is being discussed within the ethics literature, usually focusing on its implications for different questions of justice, ranging from participatory, distributive, to historical, post-colonial and feminist notions (Preston, 2013; Buck et al., 2014; Horton and Keith, 2016; Whyte, 2018). These discussions do not focus on the moral impact in the sense of techno-moral change that SAI could have on society.

One exception is Stephen Gardiner's argument for moral corruption, recently reiterated and refined by him, and Augustin Fragnière (Gardiner, 2010; Fragnière and Gardiner, 2016). Originally the argument was aimed at the claim that current generations have a responsibility to research SAI for future generations to have access to it. Gardiner criticizes this, claiming that even researching a technology such as SAI could have a troubling impact on a shared sense of responsibility for dealing with climate change. By putting resources into the research of a highly risky and uncertain technology, current generations might consider their fair share of climate action achieved, arming future generations with the know-how of how to use SAI in an emergency (Betz, 2012).

However, according to Gardiner, this would represent a perversion of the norm of responsibility. By framing SAI research as a means of equipping future generations with this technology, 'we – the current generation, and especially those in the affluent countries – are particularly vulnerable to moral corruption, that is, to the subversion of our moral discourse to our own ends' (Gardiner,

[42] It should be noted that some authors challenge this taxonomy (Buck *et al.*, 2014; Smith, 2018), arguing that the risk (or the gendered difference between care and domination) involved in deploying the different methods would be better suited to distinguishing between them. For the sake of brevity, this article adheres to the Royal Society's taxonomy.

2010, p.286). Although not explicit, the claim that merely researching a new technology can subvert an entire moral discourse is clearly an argument for techno-moral change, albeit in the form of techno-moral regress perhaps.[43]

While SAI potentially forces sweeping techno-moral changes of different kinds of norms of values, this article focuses on one norm, namely sustainability and the underlying value of nature. There are two reasons for this. One, sustainability is a central theme when discussing technologies that are aimed at alleviating the impacts of climate change. In contrast, continuous development, growth and innovation can be seen as one of the reasons why current industrialized societies are unsustainable. Two, sustainability is a norm that draws its justification from the necessity to account for the value of nature in the face of expanding human environmental impact. The ways in which technology might impact our perception of nature are especially stark in the case of geoengineering, while at the same time geoengineering is in some way considered to protect nature. This seemingly contradictory relationship makes sustainability and nature ideal candidates for the investigation of techno-moral change through SAI.

V.4.4 Sustainability and technology

We think about sustainability in the context of technology because of the predicament of striving for continuous growth and welfare on a planet with limited resources. With the Brundtland report in 1987, the notion of sustainable development gained traction as a necessary condition for equitable growth and well-being on a global scale (Brundtland, 1987). The necessity for sustainable development instead of merely development came on the back of the realization that continuous and unregulated economic growth would eventually overburden the earth's ecosystem (Meadows *et al.*, 2013). The idea of planetary boundaries served to highlight this predicament, pointing out that growth must respect both ecological and societal limits.

Squaring economic growth and affluence with other values, such as nature, is expressed in the I-PAT equation. The environmental impact (I) consists of the population of a given nation (P) times its affluence (A) times its technological efficiency (T) (Ehrlich and Holdren, 1971;

[43] Whether this fear of moral corruption is justified is difficult to answer, and requires both empirical and conceptual work. It should be noted, however, that surveys regarding public awareness of solar geoengineering tend to increase individual commitment to climate action, rather than decrease it (Merk *et al*., 2016; Cherry *et al*., 2021). In a recent book, Gernot Wagner also argues that researching solar geoengineering could be a good way to scare people into more stringent climate policies (Wagner, 2021).

Holdren and Ehrlich, 1974). Since decreasing populations or affluence seem political non-starters, reducing environmental impact is possible only through technological innovation.[44] This, in turn, brings innovation and growth into an interesting relationship with sustainability, challenging the traditional opposition between innovation and de-growth (Pesch, 2018). The paradigm of economic growth often describes sustainability as an innovation challenge. From this perspective, the solution to increasingly unsustainable ways of life is not a change in behaviour or economic systems (e.g., capitalism). Rather, sustainability is to be achieved through technological innovation, which is coupled with, instead of juxtaposed to, economic growth.

In contrast, paradigms of de-growth challenge the ideal of continuous economic growth. While growth is often coupled with innovation, one can also imagine how innovation within the ICT sector, such as in decentralized forms of communication, can lead to effective de-growth and an increase in sustainability. The seemingly paradoxical relationship between growth and de-growth becomes clear when technological development is tied to the notion of reduction and decentralization rather than expansion (Pesch, 2018). Based on this, one can see how technological development becomes an important instrument to realize sustainability under current circumstances. The connection between technology and sustainability can thus be summarized in terms of being both complex and necessary. While the quest for never-ending economic growth can be considered to be a bane to sustainability, growth that enables sustainable innovation might prove vital in dealing with certain unsustainable practices. At the same time, technological development changes what can be considered sustainable, by changing our perception of nature.

V.4.5 Nature, sustainability and SAI

When researchers propose ways of geoengineering, they usually do so to protect both human beings and the environment, or nature, from climate change. This implies that to be more sustainable, humanity should increase its impact on nature, rather than step back from it. This tacit endorsement of human control over nature is paramount to understanding the relationship between nature, sustainability and SAI. Taking control over nature is argued to be a shift from

[44] Apart from the political reasons, there is another argument for neglecting the factor of population control when accounting for environmental impact, as can be seen in the immense inequalities in terms of who emits how much. When tied to wealth, the richest 10% of the planet was responsible for close to 50% of annual emissions in 2015 (Kartha *et al.*, 2020). Also, note that both Holdren and Ehrlich acknowledge that population growth and increasing affluence are a central issue for environmental degradation (see, e.g., Holdren, 2018).

past to modern technologies by such philosophers as Martin Heidegger (1977) and Hans Jonas (1984).

> Unspoken, but self-evident for those times [antiquity], is the pervading knowledge behind it all that, for all his boundless resourcefulness, man [sic] is still small by the measure of the elements: precisely this makes his sallies into them so daring and allows those elements to tolerate his forwardness. Making free with the denizens of land and sea and air, he yet leaves the encompassing nature of those elements unchanged and their generative powers undiminished. He cannot harm them by carving out his little kingdom from theirs. They last, while his schemes have their short-lived ways. (Jonas, 1984, p.3)[45]

The first realization we have to have as modern human beings is that we no longer leave the encompassing nature of land, sea and air unchanged. Instead, by carving out our little kingdoms, humanity has had an infamous impact on the planet, most prominently (though not exclusively) through climate change.[46] Furthermore, and most pertinent in terms of sustainability, our schemes seem to last and thus affect the plant, animal and natural kingdoms even more violently.

In this view, nature is the clearly distinguishable other, separate from culture and humanity (cf. Vogel, 2015; Latour, 2017). What makes modern technology, and specifically geoengineering through SAI, particularly contentious is its active and intentional intervention in the climate system, breaking the boundary between the natural and the artificial. Although climate change has been considered an experiment of unprecedented scale itself (Revelle and Suess, 1957), there is a fundamental difference from what SAI entails, the intentional making of a new climate. By deploying SAI, '[t]he climate would become artificial in the literal sense of becoming an artefact – a product of human endeavour' (Hulme, 2014, p.95). This concept of 'arteficiation' has significant consequences for the value of nature in terms of sustainability.

On the one hand, the arteficiation of the climate seems to be an intuitively hubristic and thus unacceptable undertaking. The risks are simply too high, and the disregard for nature and the environment too brazen. From this perspective, SAI represents the final frontier of the eco-modernist endeavour, disregarding nature and anything non-human unless it can serve anthropocentric ends. Whatever is deemed sustainable is so because it upholds and furthers human interests. Nature is to be dominated through technology and must bend to humanity's will. The discourse of domination is at the very least echoed in the idea of SAI building on imaginaries of Cold War era power fantasies (Stilgoe, 2015). It is further reminiscent of mechanistic conceptions of nature, wherein domination of men over nature is central to any scientific

[45]Jonas is not referring to geoengineering, but rather gene-editing and the destructive force of nuclear power.

[46]Though not unconnected, the influence modern human beings have had on the nitrogen cycle can be considered equally impactful (Jenkinson, 2001)

undertaking. In this regard, SAI also calls for an ecofeminist perspective that is able to take into account the deeply gendered aspects of its science and discourse (Buck *et al*., 2014).

On the other hand, the reality of climate change makes denying human beings' prominent role in shaping the planet hard to dismiss. Taking the responsibility that comes with this influence seriously leads to the idea of planetary stewardship, accepting responsibility for the power humanity has accrued (Steffen *et al.*, 2011). Leslie Thiele (2019) highlights this idea when trying to bridge the gap between the staunch opponents of SAI ('Gaians') and tentative proponents ('Prometheans') through the concept of sustainability. Thiele (2019, p.475) argues that SAI could be perceived as a 'regrettable but necessary means for our species to regain the status of nature's ally, rather than go down swinging as her arrogant opponent'. Whether as Gaians or Prometheans, SAI as a technology forces us to be clear about what can and cannot be sustainable, and how we position ourselves towards nature, our environment and the non-human. We can identify this as a form of techno-moral learning (Swierstra, 2013), in that there needs to be a societal re-evaluation of what nature means when its very existence is put into question through new and emerging technologies. More specifically, it forces us to ask what we can consider sustainable, and what sustainability entails.

V.4.6 Two techno-moral vignettes

So far we have established a connection between technology and sustainability, how technology is relevant to achieving sustainability, and how both relate to the value of nature. In order to analyse SAI on the basis of techno-moral change, this section explores how the research and potential deployment of SAI can change the norm of sustainability and the value of nature. To give an outline of how SAI could shift our understanding of a sustainable future, it is useful to distinguish between two kinds of sustainability: entrenchment and transformation. Sustainability as entrenchment is the idea of continuing on the current trajectory of the eco-modernist project, focusing on finding an 'ecological balance of modern societies by means of technological and social innovation' (Adloff and Neckel, 2019, p.1018). This is entrenchment since there are no fundamental changes in how a society is organized or in the values and norms it exhibits. Rather, basic market systems (such as capitalism and consumer societies) are taken for granted, while certain elements are tweaked in a piecemeal fashion to deal with arising issues.[47] To illustrate this distinction, I engage in a type of downscaled scenario building in the form of two short techno-moral vignettes from some point in the future (Keulartz *et al*., 2004; Boenink *et al.*, 2010)

[47] This approach is reminiscent of Karl Popper's outlook in *Open Society* (Popper, 2008).

after SAI has been deployed. The two vignettes should be understood as diametrically opposed extreme cases of TMC.

Scenario one: entrenchment

It's been ten years since the first fleet of *Veilmakers* (Morton, 2015) lifted off. Temperature rise has been stabilized, only increasing at a prolonged rate, and CO2 levels have peaked. Media outlets, politicians and a swath of the public are celebrating the defeat of climate change. People live their lives, as usual. The EU's goals of complete decarbonization have slowed, so have the commitments by the US, India and China. Now that the easier parts of decarbonization have been achieved, any step further involves a great effort that is no longer economically feasible since it is easier simply to put more aerosols into the stratosphere. Those calling for structural changes are framed as climate alarmists with hidden ideological agendas – after all, technology found a way out for us! One can only imagine the pointless suffering had society made all those changes in 2021. Instead, Jeff Bezos and Elon Musk have brought together a group of billionaires who vowed to keep up the veil, assuring the general public that the termination shock was always just an unreasonable scare tactic.

This first vignette is an imaginary representation of the potential moral corruption that SAI might cause (Gardiner, 2010), taken to a polemic extreme. Opponents frequently use this possibility of entrenchment to argue against SAI (Hulme, 2014; Schneider, 2019; Dryzek and Pickering, 2019). Resistance towards SAI often entails opposition to current global power structures and economic systems. Arguing in favour of SAI is akin to defending a failed status quo (Schneider, 2019). Similarly, John Dryzek and Jonathan Pickering dismiss SAI on the basis that the infrastructure and institutions necessary to support the technology would have to be 'global, paramount, and permanent', foreclosing any other future developments (Dryzek and Pickering, 2019, p.53). The apparent impossibility of 'coming back' from a geoengineered climate makes it all the more dangerous.

Frank Adloff and Iris Hilbrich add the concept of 'control' to discuss geoengineering and concepts of sustainability (Adloff and Hilbrich, 2021). They conceptualize geoengineering as a way to enact control over nature, as the eco-modernists seek to solve pressing issues of climate change and societal problems through technological innovation (Adloff and Hilbrich, 2021, p.177). This view of geoengineering can be expanded through Dale Jamieson's argument that human beings value nature because of its autonomy: '[W]hat we value in nature is that she "does her own thing" and is largely indifferent to us' (Jamieson, 2008, p.166). Following this line of argument, one might conclude that SAI is the negation of this autonomy, enacting control rather than allowing for autonomy. From this we can summarize that control over the climate necessarily entails eliminating nature's autonomy, which is what is valuable about nature in the first place. The norm of sustainability in times of geoengineering through SAI seems to lose all meaning, if it is to stay at all connected to nature.

However, a nuanced approach can put SAI into a more positive light in the form of planetary stewardship (Steffen *et al.*, 2011). Realizing the lasting impact human societies have on their surroundings can also entail realizing a special kind of responsibility to deal with these impacts. As Steffen and colleagues point out, 'We are the first generation with the knowledge of how our activities influence the Earth System, and thus the first generation with the power and the responsibility to change our relationship with the planet' (Steffen *et al.*, 2011, p.749). From this perspective, thinking about SAI as a partial, incremental step towards dealing with and perhaps reducing the negative human impacts becomes morally necessary. In addition, the fact that such technologies are being considered could have a transformative effect on current societal understanding of nature and consequently the norm of sustainability. It could serve as a techno-moral catalyst for moral learning. The second scenario explores this.

Scenario two: transformation

It's been ten years since the first fleet of *Veilmakers* (Morton, 2015) lifted off. Temperature rise has been stabilized, only increasing at a prolonged rate, and CO2 levels have peaked. Despite the apparent victory, one can sense a reluctance for celebration. 'The veil can only keep us safe for so long – we need to act now!' reads a famous newspaper's headline. Alongside the decision to deploy SAI, most nations on the planet have agreed to increase their efforts on not just complete decarbonization, but the active reduction of atmospheric CO2. Since the veil has been up, there has been an increased public push towards radical mitigation. 'We are just suffocating ourselves if we don't follow this up with carbon sequestration', a leading climate scientist is quoted as saying. 'This is just the beginning of our shared effort'. The reality of geoengineering has also made people sensitive to broader societal issues, especially income and wealth inequalities within and between nations. The veil has made people become more aware of the climate as a shared, living and breathing space. Many countries are working on implementing deep-cutting wealth taxes, while private jets and yachts have been banned. Obviously, there is no place for this absurd luxury on a planet that is undergoing chemotherapy.

Transformation represents a future wherein deployment has underscored the necessity to become more sustainable, enforcing rather than undermining sustainability. Here, SAI and its impact on sustainability are considered to be a catalyst for societal moral change. At heart, transformation entails broad-sweeping societal changes, including behavioural and institutional adaptions. The necessity for these changes comes from a shared understanding that 'the natural and social foundations of life on earth will not be protected by means of a further economisation of sustainability' (Adloff and Neckel, 2019, p.1020)

Proponents of SAI seem not to be explicit about the question of entrenchment/transformation, although virtually all scholars talking about SAI emphasize the inevitable need to move towards a carbon-zero future. From their perspective, doing the research is a necessary precursor to learning whether it should be done or not (Wagner, 2021; cf. Hamilton, 2013b). One of the most prominent researchers on SAI, the physicist David Keith, further argues that deployment could be undertaken incrementally and only to the extent that it offsets half of the increased global warming. This would allow for monitoring progress, aborting deployment if necessary and thereby facing less risk and potential damage. Secondly, it would give a strong incentive to continue with radical mitigation (Keith, 2013, p.23).

Critics will be quick to point out that such an application of SAI is unrealistic and that Machiavellian instincts will have politicians rather increase rates of aerosols sprayed than carbon dioxide reduced or mitigated (Hamilton, 2013b). For them, SAI is a technological instrument that would further entrench existing injustices and undermine the norm of sustainability. However, this conclusion is neither logically necessary nor empirically obvious, but rather based on a pessimistic projection of how the future will develop. Transformation is an equally plausible outcome of entrenchment. However, different, implicit ways of determining what sustainability is have been reached. While entrenchment clearly sees sustainability and its accompanying values as a way of continuing current power structures, transformation leads to a holistic conceptualization of sustainability, whereby the norm becomes a necessity to uphold, protecting nature for its own sake.

It is important to keep in mind that these vignettes do not try to produce an accurate prediction of the future, nor do they serve as full-blown scenarios to be evaluated (although they could serve as a starting point for such scenario building). Rather, they serve two purposes relevant for the anticipation of societal change through new technologies. First, these vignettes of techno-moral change contextualize norms and values and show how they do not exist in a vacuum (Swierstra, 2013). Second, they encourage ethical reflection by stimulating our 'moral imagination' (Swierstra *et al*., 2009). In an attempt to realize and possibly overcome moral biases based on current standards, such a form of moral imagination allows for an anticipatory view of how society might develop through a specific instance of techno-moral change (TMC). At the same time, it forces us to confront current trajectories from a self-reflective angle: is the path we chose a good one, based on the possible futures we face?

V.4.7 How do we assess techno-moral change?

Taking a step back, one can ask after the preceding discussion what techno-moral change as a theoretical tool can tell us about the relationship between SAI and sustainability or, put differently, how a new (and physically non-existent) technology impacts a shared understanding of what ought to be of value and how to realize this value. To assess this, recall Swierstra's (2013) interpretation of morality as a force field wherein interrelated norms and values compete for societal implementation.

Entrenchment and transformation represent two different force fields of morality. Entrenchment takes the status quo for granted. There is no reason to implement fundamental change. Humanity's environmental impact is simply something to be taken into account for anthropocentric ends and does not change how we view nature. Nature is a thing to be handled for the sake of humanity and, as long as it serves human ends, any action is permissible. What is sustainable is defined by how far humans can go in using natural resources in order to carve out their empires.

In contrast, transformation represents a more radical shift from current norms and values, changing from a business-as-usual instrumental and exploitative view of nature to a caring and self-reflected perspective. The change in how nature is viewed in transformation also impacts the idea of what is or is not sustainable. While SAI might sustain our way of life for now, more is necessary (e.g., implementing a wealth-tax, redistribution of resources, political and institutional changes) in order to act sustainably. Sustainability becomes a holistic framework for evaluating not only our use of nature as a resource, but the value of nature itself.

With this in mind, there are two lessons we can draw from the process of techno-moral change through SAI. First, TMC highlights the malleability of norms and values through new societal and technological contexts. The case of SAI forces a societal confrontation of what the norm of sustainability and the value of nature mean, and whether that norm is to be followed, and that value to be realized. New technology allows (or necessitates) new kinds of behaviour, and societal norms are often quick to adapt. In this sense, technologies can tip scales in favour of certain moralities. However, it is not always clear *ex ante* which morality will end up benefiting from the impact. As the two vignettes exemplify, the impact could go either way in terms of how the norm of sustainability can be understood.

What we can take from this discussion is that TMC and the process of techno-moral learning do not necessitate the actual implementation of a new technology. SAI and geoengineering in general are very much imaginary technologies (Stilgoe, 2015). SAI exists in computer models and their empirical feasibility is drawn from inferential evidence through volcanic eruptions. Yet there is still the possibility of drawing moral lessons from the fact that these technologies are being considered in the first place, and that potential future scenarios include their deployment. The case explored here forces us to think about sustainability through technology in an age of unprecedented human impact on nature. While evaluating these potential futures of sustainability lies outside the scope of this article, the discussion has illustrated that TMC and techno-moral learning are not just means to project the future, but also mirrors of current societal and technological trajectories.

V.4.8 Conclusion

This article aims to explore techno-moral change through SAI, with a focus on how technology impacts the norm of sustainability and its accompanying value, nature. It seeks to connect technological development with the increasing necessity to act sustainably. It also tries to focus on the connection between how different ways of perceiving nature influence our understanding of sustainability and vice versa, highlighting the interdependent relationship between the norm of sustainability and the value of nature. To explore how conceptions of sustainability and nature might change through SAI, it presented two diametrically opposed views of SAI's techno-moral impact. Exploring this impact has shown two things. First, techno-moral change can go either way; it could either undermine or underscore the necessity of a sustainable future. Second, a technology does not have to be developed to have this impact. That it is being discussed as a potential future means of dealing with climate change might be enough to force techno-moral change.

VI.I Conclusion

The aim of this thesis was to explore the moral implications of SAI research, and to provide possible tools for its ethical governance. The investigation was centered around the following research question and corresponding subquestions:

> RQ: *In what ways should the desirability of researching Stratospheric Aerosol Injection be assessed, given its potentially highly disruptive effect on institutions, societal values and norms, as well as the physical environment??*
>
> SQ 1: *What are the ethically pertinent uncertainties that arise in the context of an SAI research program, and how can they be accounted for?*
>
> SQ 2: *What kind of risks in the form of disruption might the introduction of a coordinated SAI research program incur on a systemic level, and how can they be ethically governed?*
>
> SQ 3: *How can we assess the impact that an SAI research program might have on societal structures and institutions?*
>
> SQ 4: *How might an SAI research program affect complex societal values, such as sustainability?*

The concluding chapter presents the central findings of each of these questions. It then outlines some limitations of the investigation and describes further applicability beyond SAI research governance frameworks. Finally, it ties the investigation to the broader research program on the Ethics of Socially Disruptive Technologies, under the guidance of which this thesis was produced.

VI.I.I Subquestion 1: What are the ethically pertinent uncertainties that arise in the context of an SAI research program, and how can they be accounted for?

A central hurdle for the ethical evaluation of any potential SAI research program is its many risks and uncertainties. In the case of SAI, this is particularly challenging given that the arising uncertainties stem from a lack of empirical information that fundamentally undermines the political and ultimately ethical decision-making process. Uncertainty arises in the form of a lack of empirical information, alongside the ontic uncertainty stemming from inherently chaotic systems such as the climate system, wherein linear causality is not applicable. Even if it were possible to collect all the relevant data and compute it accordingly, the system's inherent complexity would prohibit exact predictions. Finally, the moral evaluation of an inherently uncertain action is itself a challenging undertaking. If there are no clear facts about the context of a given situation, the moral judgement about what is the right thing to do becomes blurry and contestable at every moment. This is compounded by the fact that an action with global implications, such as SAI research, must account for a wide array of differing, potentially mutually exclusive value-claims.

The first chapter explores this challenge by introducing the concept of normative uncertainty. Building on previous notions of moral and empirical uncertainty, it conceptualizes normative uncertainty in the following way.

> Normative uncertainties are irreducible knowledge-constraints that arise in concrete, practical decision-situations. These decision-situations are multi-lateral and empirically ambiguous, wherein the decision-makers need to reconcile a plurality of values. (Hofbauer 2023, 8)

This conception of normative uncertainties builds on existing notions from metaethical and decision-theoretical scholarship, as well as work done in the philosophy and ethics of technology. The chapter's main take-away is that normative uncertainties arising in the context of SAI research require robust and ethically reflected forms of participation and iterative governance. Participatory approaches such as mini-publics could be potential tools for the exploration of what kinds of values (i.e. concerns, needs, interests) ought to be part of a debate on SAI research governance, and what values ought to be excluded. At the same time, adaptive governance approaches provide the necessary flexibility and continuous reflection that is required when dealing with normative uncertainties.

VI.I.II Subquestion 2: What kind of risks and disruptions might a coordinated SAI research program incur on a systemic level, and how can they be ethically governed?

The second chapter takes a closer look at the kind of risks that arise in the context of SAI research. Specifically, the chapter looks at SAI research governance risks through the lens of systemic risks. Systemic risks, in contrast with more traditional forms of risk, arise in the context of complex, interrelated systems. Systemic risks are risks that potentially undermine the functioning of vital systems, such as the healthcare system, essential infrastructure, or critical supply chains. The crash of the financial markets 2008 or the Covid-19 pandemic are all examples of materialized systemic risks that have threatened the functioning of vital societal systems. Similarly, the consequences of climate change and the continuous human-induced environmental destruction represent forms of systemic risks to existing modes of living in industrialized societies.

The systemic risk approach is extremely valuable when it comes to the governance of the unintended societal side-effects of SAI research. For example, some experts argue that the introduction of a coordinated SAI research project could lead to a form of "mitigation-deterrence" (Markusson 2016), i.e., a reduction in societal and political willingness to radically mitigate, since SAI would be (mistakenly) presented as a solution to climate change. The second chapter uses the systemic risk approach as a framework to mitigate and govern this risk by anticipating the possible "weak points" that the relevant system exhibits, e.g., the energy system and its accompanying sociotechnical interdependencies.

While useful, the systemic risk approach can be enhanced through a more robust reflection on its underlying ethical assumptions about the societal implications of NESTs and its research process. Accordingly, the chapter expands on the systemic risk framework supplementing it with two central insights from the ethics of technology literature. It provides a theoretically grounded conception of participation, moving beyond mere epistemic reasons for the inclusion of a plurality of voices towards an approach that also accounts for the underlying justice reasons (based on Fraser & Honneth 2003). It further adds another level of reflection to systemic risks, highlighting the fact that the technology-society nexus are co-dependent and reciprocal spheres. This entails that the same way technology may impact society, society also drives, and influences technology through its guiding norms, values, and imaginaries.

Consequently, the second chapter seeks to supplement the systemic risk governance framework with insights from technology ethics. Concretely, answering the second subquestion introduces an ethically reflective conception of societal value dynamism and stakeholder engagement and participation, to existing systemic risk governance approaches.

VI.I.III Subquestion 3: How can we assess the impact that an SAI research program might have on societal structures and institutions?

The first two chapters assess the risks and uncertainties related to SAI research governance, with a particular focus on their ethical implications. The third chapter with its accompanying subquestion builds on these investigations, further seeking to investigate the broader context of how the interdependencies between NESTs and society impact institutions and societal structures in two central, seemingly opposing, ways. Namely, the chapter explores how an SAI research program could be disruptive, undermining existing institutions and ways of seeing the world or entrenching existing unsustainable forms of behavior.

On the one hand, SAI research is often said to be unjustifiably risky given its unprecedented global scope and its grave implications for current and future generations. Researching such a technology is an unsustainable disruption toward the energy transition and the move toward a more ecologically aware institutions and structures. On the other hand, an SAI research program is sometimes argued to be sure-fire way of eventually deploying the technology, due the socio-political dynamics of "lock-in", wherein institutional and political forces inevitably create circumstances wherein SAI will be deployed, no matter the outcome of the research project itself.

The third chapter structures these arguments around the concept of reversibility. Established approaches within the ethical and social science debates surrounding innovation, such as Responsible Research and Innovation, Value Sensitive Design, and most prominently the Precautionary Principle, all seek to account for and anticipate a new technology's potentially disruptive impacts. The main assumption of the frameworks is that researching NESTs should not come at the cost of societal control, which entails that it should be possible to reverse the impacts they had so far. The Precautionary Principle specifically has been actively incorporated by scholars such as Lauren Hartzell-Nichols (2012; 2017) to identify the potential dangers of SAI and its research. SAI deployment could itself lead to catastrophic outcomes, which disqualifies it under a certain interpretation of the Precautionary Principle, to be an acceptable means of dealing with the catastrophic impacts of climate change (Hatzell-Nichols 2017).

A question that remains underexplored in the literature however is to what degree merely researching a NEST should be evaluated in the context of these ethical frameworks. This is the main point of inquiry for the third chapter. The paper accordingly argues that while reversibility seems to be a prudent criterion to have for the development of NESTs, the extent and subject of the reversibility needs to be clearly laid out beforehand. For example, while the physical impacts of a small-scale real world research project on the scattering of stratospheric aerosols could be considered reasonably reversible, its accompanying socio-political implications may not be. The emerging claim then is that while a useful tool for the investigation of new technologies, reversibility needs to be strictly qualified and contextualized, if it is to serve as a criterion for ethical and practical guidance.

VI.I.IV Subquestion 4: How might an SAI research program affect complex societal values, such as sustainability?

The fourth and final chapter of the thesis is an exercise in techno-moral imagination (Boenink et al 2010; Swierstra 2013). Specifically, it explores how discussions on sustainability regarding its ethical and societal conceptualizations, may be transformed, disrupted, or otherwise challenged through the implementation of an SAI research program. The two central insights of this exploration are, one, showing how norms such as justice, well-being, and sustainability are interrelated and a shift in one of them will affect the conceptualization of the others. Two, societal values and norms may be disrupted, changed, or entrenched through NESTs without those technologies physically existing. It is enough for the technology to merely exist as an idea, or a “vanguard vision” by a group of experts (Hilgartner 2015). Exploring and playing with techno-moral imaginations of the future opens up a further line of debate for scenario-building workshops and any form of stakeholder engagement, and thus give a more holistic view of the kinds of societal values and concepts that are at stake when considering SAI research.

Through the introduction of two techno-moral vignettes, the chapter investigates two diametrically opposed ways in which SAI research may influence the societal conceptualization of nature. The first vignette describes a techno-fixed, neo-liberal dystopia, wherein SAI research leads to the abandonment of nature as something to be valued in and of itself. As an apparent justification and enabler for the continuous exploitation of natural resources and the non-human world, SAI research could further underscore humanity’s self-understanding of being all-powerful in controlling the environment. Solar climate engineering would fall into the hands a wealthy elite, declaring the climate crisis “resolved”, and sustainability as a value would be nothing more than a hollow label that is unable to account for deeper issues of social justice and human well-being.

In contrast, the second vignette maps out ways in which SAI research could help highlight the need for fundamental societal transformation alongside technological solutions. As with any instance of techno-moral change, SAI research would not be the sole factor for such a transformation, but rather an additional call to unambiguously accept the urgency for radical climate action. With the realization that "it has come to this" (Sapinski et al. 2020), i.e., that humanity has impacted the environment to a degree that something like the deployment of SAI is actually being considered, the societal understanding of the climate crisis deepens. Further, the fact that once humanity commits to the deployment of SAI, radical mitigation and negative emissions become even more urgent, provides a new sense of solidarity and a shared vision of a sustainable future that takes seriously how justice and well-being ought to be front and center to any climate solution.

Importantly, these vignettes are not supposed to represent realistic future outcomes, nor do they seek to anticipate techno-moral change in the real world. Instead, techno-moral imagination serves as a framework to ask morally relevant questions with regard to the consideration, development, and possible deployment of NESTs such as SAI. This in turn, makes the participatory and engagement processes surrounding SAI research more aware of the ethical and philosophical implications such a research project could carry.

VI.I.V RQ In what ways should the desirability of researching Stratospheric Aerosol Injection be assessed, given its potentially highly disruptive effect on institutions, societal values and norms, as well as the physical environment?

Based on these findings, I argue that researching SAI can be done in an ethically responsible and reflected way, if the research process is also guided by the tools and frameworks described above. Centrally, this argument goes against the assumption that researching SAI is inherently morally wrong, as it inevitably perpetuates structural injustices, hubris, or political inertia toward climate action. While SAI represents a potentially socially disruptive technology, it would be overly hasty and deterministic to pre-emptively assess its research process as being either "good" or "bad", simpliciter. Instead, as this thesis argues, SAI research should be evaluated based on various procedural tools, that actively take into account the normative uncertainties and risks involved in researching SAI, which in turn allows for an adaptive and reflexive evaluation process. By highlighting and actively incorporating the risks and uncertainties involved in SAI research, this fundamentally procedural approach for the assessment and evaluation of SAI research escapes dichotomous, static, and deterministic categorization. Similar to understanding new technologies as societal experiments (van de Poel 2013), approaching SAI research from the perspective of normative uncertainties emphasizes the malleability and dynamism of societal values, norms, and institutions and their interdependencies in relation to technological development.

Answering the overarching research question, understanding and ethically accounting for the disruptions, risks and uncertainties that solar climate engineering research invokes requires a variety of approaches and frameworks. The approaches employed throughout this thesis have revealed that uncertainty plays a central role surrounding the decision-making and evaluation of SAI research. While uncertainty poses a challenge in some regards, i.e., the physical and ethical impacts of SAI research, it can also serve as a chance for the exploration and integration of new and emerging values. Importantly, an SAI research process needs to be able to account for the value dynamism that forms part of and emerges from such a potential research process.

The exploration of normative uncertainties, and societal as well as ethical disruption through a potential SAI research program leads to the insight that SAI and potentially any other NEST needs to be evaluated in the context of a socio-technical value dynamism. This claim is founded on the assumption that societal values should be seen as dynamic and potentially shifting manifestations of interests, preferences, and reasons for (political) action, rather than being static and absolute.

On the one hand, this dynamism is a product of liberal democracies, wherein value-pluralism forms a central tenet of political life. Democracies are designed to allow for the continuous reflection, change and potential revision of existing values, all within the boundaries of certain structures that require more fundamental revisions if they are to be changed, such as constitutional changes for example. On the other hand, technological development and the pursuit of scientific endeavours is itself driven by those dynamic societal values. This is wherein the danger of perpetuating existing injustices through the research of SAI lies, i.e., that unjust existing power structures are reproduced through the research process of SAI, by ignoring or actively harming marginalized communities for example.

However, the research process also affects existing societal values and their institutional settings. In line with the philosophy of technology scholarship, a central background assumption for this thesis is that technologies and sciences shape the way we see and interpret the world. Importantly, even if a given technology is merely an imaginary concept that may never be feasible (such as SAI, for the moment at least), its very idea represents, reiterates, or undermines existing world views and values. In this sense, SAI and other NESTs may be disruptive of existing conceptions of *Nature* without necessarily being deployed (Hopster et al. 2023) For instance, learning more about how SAI can and cannot be regionally governed due to its physical impacts may change how we ethically assess its governance process, what can be considered democratic governance, and whether it is always adequate. This goes to say that the contextual setting matters for how and why certain values are applicable for the evaluation of a given situation, and that those values themselves are dynamic rather than static.

For this reason, SAI research leads to normative uncertainty within the policy and governance decision space. To account for this uncertainty from an ethical perspective, SAI research requires participatory accounts of justice, as well as anticipatory governance to continuously assess and account for dynamic societal values. At the same time, value dynamism and normative uncertainty cannot be used as an excuse to perpetuate existing injustices or justify immoral choices or values. Thus, guardrails in the form of substantive justice claims, such as justice as recognition, need to form part of the research process as well. A clear way of doing this is for communities that have been marginalized, or who's voices have been ignored to form an integral part of the research process, rather than an addition to it. A global SAI research effort for example, should be led by those most vulnerable to climate change.

VI.II. Further Applicability and Limitations

An important baseline throughout this dissertation project is the fact that interdisciplinary thinking is paramount when one seeks to understand the complexities and intricacies of governing solar climate engineering research. Thus, while guided by an argumentative and discursive analysis common to philosophical and ethical reasoning, the main insights provided by this thesis should be seen as a general means of how to assess and evaluate NESTs. In this sense, the tools provided can be of generic use when it comes to the analysis of any NESTs' societal implications, before and throughout their research phases. For example, understanding the conceptual disruption of biomimetic technologies, or cell farms that produce lab-grown meat, can help evaluate the ethical desirability of these technologies (Hopster et al. 2023). Thus, in line with the challenge of governing socially disruptive technologies (Hopster 2021; Nickel 2020), the application of ethical and philosophical analysis of technology can help enhance existing social science and governance frameworks that find use beyond the case study of SAI. The concept of normative uncertainty developed here can further be used to support the management of systemic risks (Renn et al. 2022), adaptive governance approaches, and decision-making procedures that need to account for risk and deep uncertainty (Marchau et al. 2019).

The approaches explored here can support answering the question whether SAI research is justifiable in the first place through robust and meaningful participation outlined above. However, while the tools provided are helpful in setting up and reflecting on governance processes, they need to be supported by an array of other sciences, ranging from economics to physical and climate sciences, biological and environmental research, and others in order to serve the purpose of an in-depth policy proposal. Accordingly, the proposed toolkit provides space for ethical reflection and debate on how to best tackle the governance challenges of researching NESTs such as SAI. Nevertheless, given the contextual sensitivity of this toolkit to changing values and factual circumstances, it needs to be accompanied by a number of equally essential scientific and research disciplines.

Epilogue: Reflections on The Power of Science and Technology

Modern, liberal societies are fundamentally influenced by the Enlightenment promise of continuous progress through scientific inquiry and technological innovation. The underlying allure of this promise is as enticing as it is dangerous: Any societal challenge can be resolved through the all-powerful scientific method, and its application in the form of technology. This idea is enticing, by conveying an incredible amount of autonomy and power to humanity and underscoring an optimistic outlook into the future. At the same time, this promise harbours the dangers of hubris and arrogance – dangers that stories stretching from Ovid's recounting of Icarus to Mary Shelley's Frankenstein have explored and warned against. Whether or not SAI research can lead to a hopeful future, depends on whether such a program understands and takes seriously the implications it carries, and seeks to account for the ethical and societal impacts it will have. The necessity of accounting for risk and uncertainty through meaningful participation, ethical reflections on justice, and scenario-building is relevant beyond the scope of SAI research. Any NEST has the potential to be disruptive in some way or another.

References

Adloff, Frank. 2019. ‘Sustainability’ in Paul, H. (ed.) *Critical Terms in Futures Studies*, Palgrave Macmillan, Cham, Switzerland, pp.291–7.

Adloff, Frank. and Hilbrich, Iris. 2021. ‘Practices of sustainability and the enactment of their natures/cultures: ecosystem services, rights of nature, and geoengineering’, *Social Science Information*, 60, 2, pp.168–87. https://doi.org/10.1177/0539018421998947.

Adloff, Frank. and Neckel, Sighard. 2019. ‘Futures of sustainability as modernization, transformation, and control: a conceptual framework’, *Sustainability Science,* 14, 4, pp.1015–25. https://doi.org/10.1007/s11625-019-00671-2.

Anderson, Ben. 2010. ‘Preemption, precaution, preparedness: anticipatory action and future geographies’, *Progress in Human Geography*, 34, 6, pp.777–98. https://doi.org/10.1177/0309132510362600.

Anderson, Elizabeth. 2014. *Social Movements, Experiments in Living, and Moral Progress: Case Studies from Britain's Abolition of Slavery*, Lindley Lecture, University of Kansas, available at https://kuscholarworks.ku.edu/handle/1808/14787 (accessed March 2022).

Armstrong McKay, David I., Arie Staal, Jesse F. Abrams, Ricarda Winkelmann, Boris Sakschewski, Sina Loriani, Ingo Fetzer, Sarah E. Cornell, Johan Rockström, and Timothy M. Lenton. 2022. “Exceeding 1.5°C Global Warming Could Trigger Multiple Climate Tipping Points.” *Science* 377 (6611): eabn7950. https://doi.org/10.1126/science.abn7950.

Asilomar Scientific Organizing Committee (ASOC). 2010: The Asilomar Conference Recommendations on Principles for Research into Climate Engineering Techniques, Climate Institute, Washington DC, 20006.

Bächtiger, André, Dryzek, John S., Mansbridge, Jane, & Warren, Mark E. 2018. *The Oxford handbook of deliberative democracy*. Oxford University Press.

Baker, Robert. 2013. *Before Bioethics: A History of American Medical Ethics from the Colonial Period to the Bioethics Revolution*, Oxford University Press, Oxford.

Baker, Robert. 2019. *The Structure of Moral Revolutions: Studies of Changes in the Morality of Abortion, Death, and the Bioethics Revolution*. Basic Bioethics. Cambridge, MA: The MIT Press.

Baskin, Jeremy. 2019. *Geoengineering, the Anthropocene and the End of Nature*. Cham, Switzerland: Palgrave Macmillan.

Beck, Ulrich. 1992. *Risk society: Towards a new modernity*. Theory, Culture & Society. Sage Publications.

Bellamy, Rob. 2016. “A Sociotechnical Framework for Governing Climate Engineering.” *Science, Technology, & Human Values* 41 (2): 135–62. https://doi.org/10.1177/0162243915591855.

Bergen, Jan Peter. 2016. “Reversibility and Nuclear Energy Production Technologies: A Framework and Three Cases.” *Ethics, Policy & Environment*, June. http://www.tandfonline.com/doi/abs/10.1080/21550085.2016.1173281.

Bergen, Jan. 2016. “Reversible Experiments: Putting Geological Disposal to the Test.” *Science and Engineering Ethics* 22 (3): 707–33. https://doi.org/10.1007/s11948-015-9697-2.

Betz, Gregor. 2012. ‘The case for climate engineering research: an analysis of the “arm the future” argument’, *Climatic Change*, 111, 2, pp.473–85. https://doi.org/10.1007/s10584-011-0207-5.

Biermann, Frank, Jeroen Oomen, Aarti Gupta, Saleem H. Ali, Ken Conca, Maarten A. Hajer, Prakash Kashwan, et al. 2022. "Solar Geoengineering: The Case for an International Non-Use Agreement." *WIREs Climate Change* 13 (3): e754. https://doi.org/10.1002/wcc.754.

Boenink, Marianne, Tsjalling Swierstra, and Dirk Stemerding. 2010. "Anticipating the Interaction between Technology and Morality: A Scenario Study of Experimenting with Humans in Bionanotechnology." *Studies in Ethics, Law, and Technology* 4 (2). https://doi.org/10.2202/1941-6008.1098.

Böker, Marit. 2017, February. Justification, critique and deliberative legitimacy: The limits of mini-publics. *Contemporary Political Theory*, *16*(1), 19–40. https://doi.org/10.1057/cpt.2016.11

Borup, Mads, Nik Brown, Kornelia Konrad, and Harro Van Lente. 2006. "The Sociology of Expectations in Science and Technology." *Technology Analysis & Strategic Management* 18 (3–4): 285–98. https://doi.org/10.1080/09537320600777002.

Brey, Philip. 2005 'Artifacts as social agents' in Harbers, H. and Harbers, J. (eds) *Inside the Politics of Technology*, Amsterdam University Press, Amsterdam, pp.61–84.

Brey, Philip. 2018 'The strategic role of technology in a good society', *Technology in Society,* 52, C, pp.39–45. https://doi.org/10.1016/j.techsoc.2017.02.002.

Brey, Philip. 2017. "Ethics of Emerging Technologies." *The Ethics of Technology: Methods and Approaches*, March, 175–92.

Brey, Philip. 2018. "The Strategic Role of Technology in a Good Society." *Technology in Society* 52 (February): 39–45. https://doi.org/10.1016/j.techsoc.2017.02.002.

Brundtland, Gro Harlem. 1987. *Our Common Future,* World Commission on Environment and Development, United Nations, New York.

Buck, Holly Jean, Gammon, Andrea. and Preston, Christopher. 2014. 'Gender and geoengineering', *Hypatia*, 29, 3, pp.651–69. https://doi.org/10.1111/hypa.12083.

Cairns, Rose C. 2014. "Climate Geoengineering: Issues of Path-Dependence and Socio-Technical Lock-in: Climate Geoengineering Lock-In." *Wiley Interdisciplinary Reviews: Climate Change* 5 (5): 649–61. https://doi.org/10.1002/wcc.296.

Callies, Daniel Edward. 2019. "The Slippery Slope Argument against Geoengineering Research." *Journal of Applied Philosophy* 36 (4): 675–87. https://doi.org/10.1111/japp.12345.

Cherry, Todd, Kallbekken, Steffen., Kroll, Stephan, and McEvoy, David. 2021 "Does solar geoengineering crowd out climate change mitigation efforts? Evidence from a stated preference referendum on a carbon tax', *Climatic Change,* 165, 1–2. https://doi.org/10.1007/s10584-021-03009-z.

Collingridge, David. 1980. "The Dilemma of Control." *The Social Control of Technology*, 13–22.

Committee on Developing a Research Agenda and Research Governance Approaches for Climate Intervention Strategies that Reflect Sunlight to Cool Earth, Board on Atmospheric Sciences and Climate, Committee on Science, Technology, and Law, Division on Earth and Life Studies, Policy and Global Affairs, and National Academies of Sciences, Engineering, and Medicine. 2021. *Reflecting sunlight: Recommendations for solar geoengineering research and research governance*. National Academies Press. https://doi.org/10.17226/25762

Crutzen, Paul J. 2006. "Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Resolve a Policy Dilemma?" *Climatic Change* 77 (3–4): 211–20. https://doi.org/10.1007/s10584-006-9101-y.

Davies, George. 2013 ‘Appraising weak and strong sustainability: searching for a middle ground,’ *Consilience*, 1, pp.111–24.
Dobson, Andrew. 1996 ‘Environment sustainabilities: an analysis and a typology’, *Environmental Politics*, 5, 3, pp.401–28.
Doorn, Neelke, Paolo Gardoni, and Colleen Murphy. 2019. “A Multidisciplinary Definition and Evaluation of Resilience: The Role of Social Justice in Defining Resilience.” *Sustainable and Resilient Infrastructure* 4 (3): 112–23. https://doi.org/10.1080/23789689.2018.1428162.
Dryzek, John S., and Jonathan Pickering. 2019. *The Politics of the Anthropocene*. First edition. Oxford, United Kingdom ; New York: Oxford University Press.
Ehrlich, Paul. and Holdren, John. 1971 ‘Impact of population growth’, *Science*, 171, 3977, pp.1212–17.
European Commission. 2019 *The European Green Deal*, Brussels, COM 640, available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52019DC0640&from=EN (accessed March 2022).
Fishkin, James. 2009. *When the people speak: Deliberative democracy and public consultation*. OUP Oxford.
Florin, Marie-Valentine, and Anjali Nursimulu. 2018. “IRGC Guidelines for the Governance of Systemic Risks,” October. https://doi.org/10.5075/EPFL-IRGC-257279.
Florin, Marie-Valentine. 2013. “IRGC’s Approach to Emerging Risks.” *Journal of Risk Research* 16 (3–4): 315–22. https://doi.org/10.1080/13669877.2012.729517.
Florin, Marie-Valentine. 2022. “Risk Governance and ‘responsible Research and Innovation’ Can Be Mutually Supportive.” *Journal of Risk Research* 25 (8): 976–90. https://doi.org/10.1080/13669877.2019.1646311.
Foley, Rider, Guston, David, & Sarewitz, Daniel. 2018. Towards the anticipatory governance of geoengineering. In J. J. Blackstock & S. Low (Eds.), *Geoengineering our climate?* (pp. 223–243). Routledge.
Folke, Carl. 2006. “Resilience: The Emergence of a Perspective for Social–Ecological Systems Analyses.” *Global Environmental Change* 16 (3): 253–67. https://doi.org/10.1016/j.gloenvcha.2006.04.002.
Fragnière, Augustin. and Gardiner, Stephen. 2016. ‘Why geoengineering is not “Plan B”’ in Preston, C., *Climate Justice and Geoengineering*, Rowman & Littlefield, Lanham MD, pp.15–31.
Frank, Lilly. 2020. ‘What do we have to lose? Offloading through moral technologies: moral struggle and progress’, *Science Engineering Ethics*. https://doi.org/10.1007/s11948-019-00099-y.
Fraser, Nancy, and Axel Honneth. 2003. *Redistribution or Recognition? A Political-Philosophical Exchange*. London ; New York: Verso.
Friedman, Batya, Peter H. Kahn Jr., and Alan Borning. 2008. “Value Sensitive Design and Information Systems.” In *The Handbook of Information and Computer Ethics*, 69–101. https://doi.org/10.1002/9780470281819.ch4.
Frisch, Mathias. 2013. Modeling climate policies: A Critical look at integrated assessment models. *Philosophy & Technology*, *26*(2), 117–137. https://doi.org/10.1007/s13347-013-0099-6

Galaz, Victor. 2012. ‘Geo-engineering, governance, and social-ecological systems: critical issues and joint research needs’, *Ecology and Society*, 17, 1. http://www.jstor.org/stable/26269023.
Gardiner, Stephen. 2006. A core precautionary principle*. *The Journal of Political Philosophy*, *14*(1), 33–60. https://doi.org/10.1111/j.1467-9760.2006.00237.x
Gardiner, Stephen. 2009. “Is ‘Arming the Future’ with Geoengineering Really the Lesser Evil? Some Doubts About the Ethics of Intentionally Manipulating the Climate System.” SSRN Scholarly Paper ID 1357162. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=1357162.
Gardiner, Stephen. 2010. ‘Is “arming the future” with geoengineering really the lesser evil? Some doubts about the ethics of intentionally manipulating the climate system’ in Gardiner, S., Caney, S.,
Gardiner, Stephen and Fragnière, Augustin. 2018. “The Tollgate Principles for the Governance of Geoengineering: Moving Beyond the Oxford Principles to an Ethically More Robust Approach.” *Ethics, Policy & Environment* 21 (2): 143–74. https://doi.org/10.1080/21550085.2018.1509472.
Genus, Audley, and Andy Stirling. 2018. “Collingridge and the Dilemma of Control: Towards Responsible and Accountable Innovation.” *Research Policy* 47 (1): 61–69. https://doi.org/10.1016/j.respol.2017.09.012.
Giddens, Anthony. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge, GB: Polity press, 1986.
Grunwald, Armin. 2015. “Technology Assessment and Design for Values.” In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, edited by Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel, 67–86. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_4.
Guston, Dave. 2014. Understanding “Anticipatory governance”. *Social Studies of Science, 44*(2), 218–242. https://doi.org/10.1177/0306312713508669
Guston, Dave, & Sarewitz, Daniel. 2002. Real-Time technology assessment. *Technology in Society*, *24*(1–2), 93–109. https://doi.org/10.1016/S0160-791X(01)00047-1
Haasnoot, Marjolijn, Kwakkel, Jan H., Walker, Warren E., & Ter Maat, Judith. 2013. Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change*, *23*(2), 485–498. https://doi.org/10.1016/j.gloenvcha.2012.12.006
Hale, Benjamin. 2012. “The World That Would Have Been: Moral Hazard Arguments against Geoengineering.” *Engineering the Climate: The Ethics of Solar Radiation Management* 113.
Hamilton, Clive. 2013a. *Earthmasters: The Dawn of the Age of Climate Engineering*, Yale University Press, New Haven.
Hamilton, Clive. 2013b. “No, we should not just “at least do the research”, *Nature News*, 496, 7444, p.139. https://doi.org/10.1038/496139a.
Hamilton, Clive. 2014. “Ethical anxieties about geoengineering” in Sandler, R. (ed.) *Ethics and Emerging Technologies*, Palgrave Macmillan, London, pp.439–55. https://doi.org/10.1057/9781137349088_29.
Hamilton, Clive. 2014. “Geoengineering and the Politics of Science.” *Bulletin of the Atomic Scientists* 70 (3): 17–26. https://doi.org/10.1177/0096340214531173.

Hansson, Sven Ove. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. http://grail.eblib.com.au/patron/FullRecord.aspx?p=1514205.

Hansson, Sven Ove, ed. 2017. *The Ethics of Technology: Methods and Approaches*. Philosophy, Technology and Society. London ; New York: Rowman & Littlefield International, Ltd.

Hartzell-Nichols, Lauren. 2012. Precaution and solar radiation management. *Ethics, Policy & Environment*, *15*(2), 158–171. https://doi.org/10.1080/21550085.2012.685561

Hartzell-Nichols, L. 2017. *Precautionary principles, catastrophes, and climate change*. Routledge. https://doi.org/10.4324/9781315309330

Heidegger, Martin. 1977. *The Question Concerning Technology, and Other Essays*. New York: Garland Pub.

Henke, Christopher R., and Benjamin Sims. *Repairing Infrastructures: The Maintenance of Materiality and Power*. MIT Press, 2020.

Heyward, Clare, and Steve Rayner. “A Curious Asymmetry: Social Science Expertise and Geoengineering.” *Climate Geoengineering Governance Project Working Paper* 7 (2013).

Hilgartner, Stephen. 2015. “Capturing the Imaginary. Vanguards, Visions and the Synthetic Biology Revolution.” In *Science and Democracy. Making Knowledge and Making Power in the Biosciences and Beyond*, edited by Stephen Hilgartner, A. Clark Miller, and Rob Hagendijk, 33–55. New York: Routledge.

Hofbauer, Benjamin. 2022. “Techno-Moral Change through Solar Geoengineering: How Geoengineering Challenges Sustainability.” *Prometheus*, June 1, 2022. https://doi.org/10.13169/prometheus.38.1.0082.

Hofbauer, Benjamin. (2023) “Normative Uncertainty in Solar Climate Engineering Research Governance.” *Ethics, Policy & Environment* 0, no. 0 (May 26, 2023): 1–20. https://doi.org/10.1080/21550085.2023.2216148.

Hofbauer, Benjamin. 2023b. “Systemic risks and solar climate engineering research. Integrating technology ethics into the governance of systemic risks”, Journal of Risk Research, DOI: 10.1080/13669877.2023.2288010

Holdren, John. 2018. “A brief history of IPAT”, *Journal of Population and Sustainability*, 2, 2, pp.66–74.

Holdren, John. and Ehrlich, Paul. 1974. “Human population and the global environment: population growth, rising per capita material consumption, and disruptive technologies have made civilization a global ecological force”, *American Scientist*, 62, 3, pp.282–92.

Hopster, Jeroen. 2021. “What Are Socially Disruptive Technologies?” *Technology in Society* 67 (November): 101750. https://doi.org/10.1016/j.techsoc.2021.101750.

Hopster, J. K. G., Gerola, A., Hofbauer, B., Löhr, G., Rijssenbeek, J., & Korenhof, P. (2023). Who owns NATURE? Conceptual appropriation in discourses on climate and biotechnologies. Environmental Values, 0(0). https://doi.org/10.1177/09632719231196535

Horton, Joshua. and Keith, David. 2016. “Solar geoengineering and obligations to the global poor” in Preston, Christopher (ed.) *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene*, Rowman and Littlefield, Lanham MD, pp.79–92.

Horton, Joshua , and Keith, David. 2016. “Solar geoengineering and obligations to the global poor” in C. Horton, Joshua, and Keith, David. 2016. “Solar Geoengineering and Obligations to the Global Poor.” Edited by Christopher J. Preston. *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene.*, 79–92.

Hourdequin, Marion. 2016. "Justice, recogntion and climate change." in Preston, Christopher J. (Ed.), *Climate justice and geoengineering: Ethics and policy in the atmospheric anthropocene* (pp. 33–48). Rowman & Littlefield.

Hourdequin, Marion. 2019. Geoengineering justice: The role of recognition. *Science, Technology, & Human Values*, *44*(3), 448–477. https://doi.org/10.1177/0162243918802893

Huesemann, Michael and Huesemann, Joyce. 2011. *Techno-Fix: Why Technology won't Save us or the Environment*, New Society Publishers, Gabriola Island, Canada.

Hulme, Mike. 2014. *Can Science Fix Climate Change: A Case Against Climate Engineering.* Hoboken: Wiley. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=795237.

Humphrey, Mathew. 2001. "Three Conceptions of Irreversibility and Environmental Ethics: Some Problems." *Environmental Politics* 10 (1): 138–54. https://doi.org/10.1080/714000510.

Intergovernmental Panel on Climate Change (IPCC). 2023. *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009157896.

Irvine, Pete, Lunt, Daniel, Stone, Emma, and Ridgwell, Andy. 2009 "The fate of the Greenland ice sheet in a geoengineered, high CO2 world", *Environmental Research Letters*, 4, 4. 045109. https://doi.org/10.1088/1748-9326/4/4/045109.

Preston, Christopher J. (Ed.). 2016. *Climate justice and geoengineering: Ethics and policy in the atmospheric anthropocene* (79–92).Rowman & Littlefield International.

Jamieson, Dale. 1996. Ethics and intentional climate change. *Climatic Change*, *33*(3), 323–336. https:// doi.org/10.1007/BF00142580

Jamieson, Dale. 2008. *Ethics and the Environment: An Introduction*, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511806186.

Jasanoff, Sheila. 2007. Technologies of humility. *Nature*, *450*(7166), 33. https://doi.org/10.1038/450033a

Jasanoff, Sheila, and Kim, Sang-Hyun (eds). 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*, University of Chicago Press, Chicago.

Jebari, Joseph, Olúfẹ́mi O. Táíwò, Talbot M. Andrews, Valentina Aquila, Brian Beckage, Mariia Belaia, Maggie Clifford, et al. 2021. "From Moral Hazard to Risk-Response Feedback." *Climate Risk Management* 33 (January): 100324. https://doi.org/10.1016/j.crm.2021.100324.

Jenkinson, David S. 2001. "The impact of humans on the nitrogen cycle, with focus on temperate arable agriculture", *Plant and Soil*, 228, 1, pp.3–15. https://doi.org/10.1023/A:1004870606003.

Jonas, Hans. 1984. *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*, University of Chicago Press, Chicago.

Jongsma, Karin Rolanda, and Martin Sand. 2017. "The Usual Suspects: Why Techno-Fixing Dementia Is Flawed." *Medicine, Health Care and Philosophy* 20 (1): 119–30. https://doi.org/10.1007/s11019-016-9747-9.

Kartha, Sivan, Kemp-Benedict, Eric, Ghosh, Emily, Nazareth, Anisha and Gore, Tim. 2020. *The Carbon Inequality Era: An Assessment of the Global Distribution of Consumption*

Emissions among Individuals from 1990 to 2015 and Beyond, research report, Stockholm Environment Institute and Oxfam International, Oxford.
Kaufman, George G., and Kenneth E. Scott. 2003. “What Is Systemic Risk, and Do Bank Regulators Retard or Contribute to It?” *The Independent Review* 7 (3): 371–91.
Keith, David. 2001. Geoengineering. *Nature*, *409*(6818), 420. https://doi.org/10.1038/35053208
Keith, David. 2013. *A Case for Climate Engineering*, MIT Press., Cambridge MA.
Keith, David, & Irvine, Pete. 2016. Solar geoengineering could substantially reduce climate risks—A research hypothesis for the next decade. *Earth’s Future*, *4*(11), 549–559. https://doi.org/10.1002/ 2016EF000465
Kemp, Luke, Chi Xu, Joanna Depledge, Kristie L. Ebi, Goodwin Gibbins, Timothy A. Kohler, Johan Rockström, et al. 2022. “Climate Endgame: Exploring Catastrophic Climate Change Scenarios.” *Proceedings of the National Academy of Sciences* 119 (34): e2108146119. https://doi.org/10.1073/pnas.2108146119.
Keulartz, Jozef, Maartje Schermer, Michiel Korthals, and Tsjalling Swierstra. 2004. “Ethics in Technological Culture: A Programmatic Proposal for a Pragmatist Approach.” *Science, Technology, & Human Values* 29 (1): 3–29. https://doi.org/10.1177/0162243903259188.
Klein, Naomi. 2015. *This Changes Everything: Capitalism Vs. The Climate*. Simon and Schuster,
Kolbert, Elizabeth. 2021. *Under a White Sky: The Nature of the Future*. First edition. New York: Crown.
Kravitz, Ben, and Douglas G. MacMartin. 2020. “Uncertainty and the Basis for Confidence in Solar Geoengineering Research.” *Nature Reviews Earth & Environment* 1 (1): 64–75. https://doi.org/10.1038/s43017-019-0004-7.
Kriebel, David, Tickner, Joel, Epstein, Paul, Lemons, John, Levins, Rrichard, Loechler, Edward L., Quinn, Margaret, Rudel, Ruthann, Schettler, Ted, & Stoto, Michael. 2001. The precautionary principle in environmental science. *Environmental Health Perspectives*, *109*(9), 871–876. https://doi.org/10.1289/ehp.01109871
Kwakkel, Jan H., & Haasnoot, Marjolijn. 2019. Supporting DMDU: A taxonomy of approaches and tools. In V. A. W. J. Marchau, W. E. Walker, P. J. T. M. Bloemen, & S. W. Popper (Eds.), *Decision making under deep uncertainty: From theory to practice* (pp. 355–374). Springer International Publishing. https:// doi.org/10.1007/978-3-030-05252-2_15
Latour, Bruno. 2017. *Facing Gaia: Eight Lectures on the New Climatic Regime,* John Wiley & Sons, London.
Latour, Bruno, and Couze Venn. 2002. “Morality and Technology.” *Theory, Culture & Society* 19, no. 5–6: 247–60.
Lenferna, Georges Alexandre, Rick D. Russotto, Amanda Tan, Stephen M. Gardiner, and Thomas P. Ackerman. 2017. “Relevant Climate Response Tests for Stratospheric Aerosol Injection: A Combined Ethical and Scientific Analysis.” *Earth’s Future* 5 (6): 577–91. https://doi.org/10.1002/2016EF000504.
Low, Sean, and Holly Jean Buck. “The Practice of Responsible Research and Innovation in ‘Climate Engineering.’” *WIREs Climate Change* 11, no. 3 (May 2020): e644. https://doi.org/10.1002/wcc.644.
MacAskill, William. 2016. Normative uncertainty as a voting problem. *Mind*, *125*(500), 967–1004. https:// doi.org/10.1093/mind/fzv169
MacAskill, William, Bykvist, Krister, & Ord, Toby (Eds.) 2020. *Moral uncertainty*. New product. Oxford University Press.

Maheshwari, Kritika, and Sven Nyholm. 2022. "Dominating Risk Impositions." *The Journal of Ethics* 26 (4): 613–37. https://doi.org/10.1007/s10892-022-09407-4.
Manson, Neil A. 2002. Formulating the precautionary principle. *Environmental Ethics*, *24*(3), 263–274. https://doi.org/10.5840/enviroethics200224315
Marchau, Vincent A. W. J., Warren E. Walker, Pieter J. T. M. Bloemen, and Steven W. Popper, eds. 2019. *Decision Making under Deep Uncertainty: From Theory to Practice*. 1st ed. 2019. Cham: Springer International Publishing : Imprint: Springer. https://doi.org/10.1007/978-3-030-05252-2.
Markusson, Nils, Duncan McLaren, and David Tyfield. 2018. "Towards a Cultural Political Economy of Mitigation Deterrence by Negative Emissions Technologies (NETs)." *Global Sustainability* 1 (January): e10. https://doi.org/10.1017/sus.2018.10.
McKinnon, Catriona. 2019. "Sleepwalking into Lock-in? Avoiding Wrongs to Future People in the Governance of Solar Radiation Management Research." *Environmental Politics* 28 (3): 441–59. https://doi.org/10.1080/09644016.2018.1450344.
McLaren, Duncan. 2016. "Mitigation Deterrence and the 'Moral Hazard' of Solar Radiation Management." *Earth's Future* 4 (12): 596–602. https://doi.org/10.1002/2016EF000445.
Meadows, Donella, Jorgen, Randers, and Meadows, Dennis. 2013. *The Limits to Growth (1972): The Future of Nature*, Yale University Press, New Haven. http://www.degruyter.com/document/ doi/10.12987/9780300188479-012/html.
Merk, Christine, Pönitzsch, Gert and Rehdanz, Katrin. 2016 "Knowledge about aerosol injection does not reduce individual mitigation efforts", *Environmental Research Letters*, 11, 5, 054009. https://doi. org/10.1088/1748-9326/11/5/054009.
Merkerk, Rutger O. van, and Harro van Lente. 2005. "Tracing Emerging Irreversibilities in Emerging Technologies: The Case of Nanotubes." *Technological Forecasting and Social Change* 72 (9): 1094–1111. https://doi.org/10.1016/j.techfore.2004.10.003.
Midgley, Mary. 2006. *Science and Poetry*. Routledge Classics. London ; New York: Routledge.
Moreno-Cruz, Juan B., & Keith, David. 2013. Climate policy under uncertainty: A case for solar geoengineering. *Climatic Change*, *121*(3), 431–444. https://doi.org/10.1007/s10584-012-0487-4
Milkoreit, Manjana, Jennifer Hodbod, Jacopo Baggio, Karina Benessaiah, Rafael Calderón-Contreras, Jonathan F Donges, Jean-Denis Mathias, Juan Carlos Rocha, Michael Schoon, and Saskia E Werners. "Defining Tipping Points for Social-Ecological Systems Scholarship—an Interdisciplinary Literature Review." *Environmental Research Letters* 13, no. 3 (March 1, 2018): 033005. https://doi.org/10.1088/1748-9326/aaaa75.
Morton, Oliver. 2015. *The Planet Remade: How Geoengineering Could Change the World*, Granta, London.
National Academies of Sciences, Engineering. 2021. *Reflecting Sunlight: Recommendations for Solar Geoengineering Research and Research Governance*. https://doi.org/10.17226/25762.
Neuber, Frederike. and Ott, Konrad. 2020. "The buying time argument within the solar radiation management discourse", *Applied Sciences*, 10, 13, paper 4637. https://doi.org/10.3390/app10134637.
Newell, Peter, Srivastava, Shilpi, Naess, Lars Otto, Torres Contreras, Gerardo, & Price, Roz. 2021. Toward transformative climate justice: An emerging research agenda. *WIREs Climate Change*, *12*(6), e733. https:// doi.org/10.1002/wcc.733

Neyrat, Frédéric. 2020. *The Unconstructable Earth: An Ecology of Separation*. Fordham University Press. https://doi.org/10.1515/9780823282609.
Nickel, Philip, Kudina, Olya, and van de Poel, Ibo. 2022. "Moral uncertainty in technomoral change: bridging the explanatory gap", *Perspectives on Science*, 18 January.
Nickel, Philip. 2020. "Disruptive Innovation and Moral Uncertainty." *NanoEthics*, October. https://doi.org/10.1007/s11569-020-00375-3.
Niemeier, Ulrike, & Tilmes, Simone. 2017. Sulfur injections for a cooler planet. *Science, 357*(6348), 246–248. https://doi.org/10.1126/science.aan3317
Niemeier, Ulrike, & Timmreck, Claudia. 2015. "What Is the Limit of Climate Engineering by Stratospheric Injection of SO_2?" *Atmospheric Chemistry and Physics* 15 (16): 9129–41. https://doi.org/10.5194/acp-15-9129-2015.
Nordmann, Alfred. "If and Then: A Critique of Speculative NanoEthics." *NanoEthics* 1, no. 1 (May 21, 2007): 31–46. https://doi.org/10.1007/s11569-007-0007-6.
Nozick, Robert. 1974. *Anarchy, state, and Utopia* (Vol. 5038). Basic Books.
Nussbaum, Martha. 1990. Aristotelian social democracy. In R. B. Douglass, G. M. Mara, & S. H. Richardson (Eds.), *Liberalism and the good* (pp. 203–243). Routledge.
Nyholm, Sven. 2018. "Attributing Agency to Automated Systems: Reflections on Human–Robot Collaborations and Responsibility-Loci." *Science and Engineering Ethics* 24 (4): 1201–19. https://doi.org/10.1007/s11948-017-9943-x.
Orlikowski, Wanda J. 1992. "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3 (3): 398–427.
Owen, Richard, Stilgoe, Jack, Macnaghten, Phil, Gorman, Mike, Fisher, Erik, & Guston, Dave. 2013. A framework for responsible innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation* (pp. 27– 50). John Wiley & Sons. https://doi.org/10.1002/9781118551424.ch2
Otto, Ilona M., Jonathan F. Donges, Roger Cremades, Avit Bhowmik, Richard J. Hewitt, Wolfgang Lucht, Johan Rockström, et al. "Social Tipping Dynamics for Stabilizing Earth's Climate by 2050." *Proceedings of the National Academy of Sciences* 117, no. 5 (February 4, 2020): 2354–65. https://doi.org/10.1073/pnas.1900577117.
Pamplany, Augustine, Bert Gordijn, and Patrick Brereton. 2020. "The Ethics of Geoengineering: A Literature Review." *Science and Engineering Ethics* 26 (6): 3069–3119. https://doi.org/10.1007/s11948-020-00258-6.
Parker, Andy, and Irvine, Pete. 2018 "The risk of termination shock from solar geoengineering", *Earth's Future,* 6, 3, pp.456–67. https://doi.org/10.1002/2017EF000735.
Pesch, Udo. 2018. "Paradigms and paradoxes: the futures of growth and degrowth", *International Journal of Sociology and Social Policy*, 38, 11/12, pp.1133–46. https://doi.org/10.1108/IJSSP-03- 2018-0035.
Poel, Ibo van de. 2016. "An Ethical Framework for Evaluating Experimental Technology." *Science and Engineering Ethics* 22 (3): 667–86. https://doi.org/10.1007/s11948-015-9724-3.
Poel, Ibo van de, Lily Eva Frank, Julia Hermann, Jeroen Hopster, Dominic Lenzi, Sven Nyholm, Behnam Taebi, and Elena Ziliotti. *Ethics of Socially Disruptive Technologies: An Introduction*. Open Book Publishers, 2023. https://doi.org/10.11647/obp.0366.
Pols, A.J.K., and H.A. Romijn. 2017. "Evaluating Irreversible Social Harms." *Policy Sciences* 50 (3): 495–518. https://doi.org/10.1007/s11077-017-9277-1.

Popper, Karl. 2008 *The Open Society and its Enemies*, vol. 2: *Hegel and Marx*, Routledge, London
Preston, Christopher J. 2013. “Ethics and Geoengineering: Reviewing the Moral Issues Raised by Solar Radiation Management and Carbon Dioxide Removal: Ethics & Geoengineering.” *Wiley Interdisciplinary Reviews: Climate Change* 4 (1): 23–37. https://doi.org/10.1002/wcc.198.
Preston, Christopher. 2016. *Climate Justice and Geoengineering: Ethics and Policy in the Atmospheric Anthropocene*, Rowman & Littlefield, Lanham MD.
Rayner, Steve, Clare Heyward, Tim Kruger, Nick Pidgeon, Catherine Redgwell, and Julian Savulescu. 2013. “The Oxford Principles.” *Climatic Change* 121 (3): 499–512. https://doi.org/10.1007/s10584-012-0675-2.
Renn, Ortwin, Manfred Laubichler, Klaus Lucas, Wolfgang Kröger, Jochen Schanze, Roland W. Scholz, and Pia-Johanna Schweizer. 2022. “Systemic Risks from Different Perspectives.” *Risk Analysis* 42 (9): 1902–20. https://doi.org/10.1111/risa.13657.
Renn, Ortwin. 2008. *Risk Governance: Coping with Uncertainty in a Complex World*. Earthscan Risk in Society Series. London ; Sterling, VA: Earthscan.
Revelle, Roger and Suess, Hans. 1957. “Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO2 during the past decades”, *Tellus*, 9, 1, pp.18–27.
Robeyns, Ingrid. 2006. The capability approach in practice. *The Journal of Political Philosophy*, *14*(3), 351–376.
Robock, Alan, Allison Marquardt, Ben Kravitz, and Georgiy Stenchikov. “Benefits, Risks, and Costs of Stratospheric Geoengineering.” *Geophysical Research Letters* 36, no. 19 (2009). https://doi.org/10.1029/2009GL039209.
Sand, Martin. 2018. *Futures, Visions, and Responsibility: An Ethics of Innovation*, Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-22684-8.
Sand, Martin, Hofbauer, Benjamin P., and Alleblas, Joost. 2023. “Techno-Fixing Non-Compliance - Geoengineering, Ideal Theory and Residual Responsibility.” *Technology in Society* 73 (May): 102236. https://doi.org/10.1016/j.techsoc.2023.102236.
Sandin, Per. 1999. Dimensions of the precautionary principle. *Human and Ecological Risk Assessment: An International Journal*, *5*(5), 889–907. https://doi.org/10.1080/10807039991289185
Schneider, Linda. 2019. “Fixing the Climate? How Geoengineering Threatens to Undermine the SDGs and Climate Justice.” *Development* 62 (1–4): 29–36. https://doi.org/10.1057/s41301-019-00211-6.
Schweizer, Pia-Johanna, and Ortwin Renn. 2019. “Governance of Systemic Risks for Disaster Prevention and Mitigation.” *Disaster Prevention and Management: An International Journal* 28 (6): 862–74. https://doi.org/10.1108/DPM-09-2019-0282.
Schweizer, Pia-Johanna. “Systemic Risks – Concepts and Challenges for Risk Governance.” *Journal of Risk Research* 24, no. 1 (January 2, 2021): 78–93. https://doi.org/10.1080/13669877.2019.1687574.
Sen, Amartya. 1987. “Equality of what?” In J. Rawls & S. M. McMurrin (Eds.), *Liberty, equality, and law: Selected tanner lectures on moral philosophy* (pp. 197–220). University of Utah Press.
Sepielli, Andrew. 2014. What to do when you don’t know what to do when you don’t know what to do. . .. *Noûs*, *48*(3), 521–544. https://doi.org/10.1111/nous.12010

Shackley, Simon, and Ken Green. 2007. "A Conceptual Framework for Exploring Transitions to Decarbonised Energy Systems in the United Kingdom." *Energy* 32 (3): 221–36. https://doi.org/10.1016/j.energy.2006.04.010.
Shackley, Simon, and Michael Thompson. 2012. "Lost in the Mix: Will the Technologies of Carbon Dioxide Capture and Storage Provide Us with a Breathing Space as We Strive to Make the Transition from Fossil Fuels to Renewables?" *Climatic Change* 110 (1–2): 101–21. https://doi.org/10.1007/s10584-011-0071-3.
Shepherd, John, Ken Caldeira, Joanna Haigh, David Keith, Brian Launder, Georgina Mace, Gordon MacKerron, et al. "Geoengineering the Climate - Science, Governance and Uncertainty." *The Royal Society*, 2009. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/publications/2009/8693.pdf.
Smith, Patrick. 2018. "Legitimacy and non-domination in solar radiation management research", *Ethics, Policy & Environment*, 21, 3, pp.341–61. https://doi.org/10.1080/21550085.2018.1562528.
Steel, Daniel. 2015. *Philosophy and the precautionary principle: Science, evidence, and environmental policy*. Cambridge University Press.
Steffen, Will, Persson, Åsa, Deutsch, Lisa, Zalasiewicz, Jan, Williams, Mark, Richardson, Katherine, Crumley, Carole *et al.* 2011. "The anthropocene: from global change to planetary stewardship", *AMBIO*, 40, 7, pp.739–61. https://doi.org/10.1007/s13280-011-0185-x.
Stilgoe, Jack. 2015. *Experiment earth: Responsible innovation in geoengineering*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/delft/detail.action?docID=1975232
Stilgoe, Jack, Owen, Richard, & Macnaghten, Phil. 2013. Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008
Stirling, Andy. 2010. Keep it complex. *Nature*, *468*(7327), 1029–1031. https://doi.org/10.1038/4681029a
Svoboda, Toby, Irvine, Pete, Callies, Daniel and Sugiyama, Masahiro. 2018. "The potential for climate engineering with stratospheric sulfate aerosol injections to reduce climate injustice", *Journal of Global Ethics*, 14, 3, pp.353–68. https://doi.org/10.1080/17449626.2018.1552180.
Swierstra, Tsjalling. 2013. "Nanotechnology and technomoral change", *Ethics & Politics*, 15, 1, pp.200–19.
Swierstra, Tsjalling and te Molder, Hedwig. 2012. "Risk and soft impacts" in Roeser, Sabine, Hillerbrand, R., Peterson, M. and Sandin, P. (eds) *Handbook of Risk Theory*, Springer, Dordrecht pp.1049–66. https://doi.org/10.1007/978-94-007-1433-5_42.
Swierstra, Tsjalling, Dirk Stemerding, and Marianne Boenink. 2009. "Exploring Techno-Moral Change: The Case of the ObesityPill." In *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments.*, edited by Paul Sollie and Marcus Düwell, 119–38. The International Library of Ethics, Law and Technology. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-2229-5_9.
Szerszynski, Bronislaw, Matthew Kearnes, Phil Macnaghten, Richard Owen, and Jack Stilgoe. 2013. "Why Solar Radiation Management Geoengineering and Democracy Won't Mix." *Environment and Planning A: Economy and Space* 45 (12): 2809–16. https://doi.org/10.1068/a45649.

Taebi, Behnam. 2017. Bridging the gap between social acceptance and ethical acceptability: Perspective. *Risk Analysis*, *37*(10), 1817–1827. https://doi.org/10.1111/risa.12734
Taebi, Behnam, Jan H. Kwakkel, and Céline Kermisch. 2020. "Governing Climate Risks in the Face of Normative Uncertainties." *WIREs Climate Change* 11 (5): e666. https://doi.org/10.1002/wcc.666.
Táíwò, Olúfẹmi O., & Talati, Shuchi. 2021. Who are the engineers? Solar geoengineering research and justice. *Global Environmental Politics*, *22*(1), 1–7. https://doi.org/10.1162/glep_a_00620
Tang, Aaron, and Luke Kemp. 2021. "A Fate Worse Than Warming? Stratospheric Aerosol Injection and Global Catastrophic Risk." *Frontiers in Climate* 3 (November): 720312. https://doi.org/10.3389/fclim.2021.720312.
Taylor, Charles. 2004. *Modern Social Imaginaries*. Public Planet Books. Durham: Duke University Press.
Thiele, Leslie. 2019. "Geoengineering and sustainability", *Environmental Politics*, 28, 3, pp.460–79. https://doi.org/10.1080/09644016.2018.1449602.
Tilmes, Simone, Douglas G. MacMartin, Jan T. M. Lenaerts, Leo van Kampenhout, Laura Muntjewerf, Lili Xia, Cheryl S. Harrison, et al. 2020. "Reaching 1.5 and 2.0 °C Global Surface Temperature Targets Using Stratospheric Aerosol Geoengineering." *Earth System Dynamics* 11 (3): 579–601. https://doi.org/10.5194/esd-11-579-2020.
Trisos, Christopher, Amatulli, Giuseppe, Gurevitch, Jessica, Robock, Alan, Xia, Lili and Zambri, Brian. 2018. "Potentially dangerous consequences for biodiversity of solar geoengineering implementation and termination", *Nature Ecology & Evolution*, 2, 3, pp. 475–82. https://doi.org/10.1038/s41559-017-0431-0.
Tuana, Nancy, Sriver, Ryan L, Svoboda, Toby, Olson, Roman, Irvine, Pete, Haqq-Misra, Jacob, & Keller, Klaus. 2012. Towards integrated ethical and scientific analysis of geoengineering: A research agenda. *Ethics, Policy & Environment*, *15*(2), 136–157. https://doi.org/10.1080/21550085.2012.685557
Uffelen, Nynke van. 2022. "Revisiting Recognition in Energy Justice." *Energy Research & Social Science* 92 (October): 102764. https://doi.org/10.1016/j.erss.2022.102764.
Véliz, Carissa. 2021. *Privacy Is Power: Why and How You Should Take Back Control of Your Data*. London: Corgi Books.
Verbeek, Peter-Paul. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago ; London: The University of Chicago Press.
Verbruggen, Aviel. 2013. "Revocability and Reversibility in Societal Decision-Making." *Ecological Economics*, New Climate Economics, 85 (January): 20–27. https://doi.org/10.1016/j.ecolecon.2012.10.011.
Vogel, Steven. 2015. *Thinking like a Mall: Environmental Philosophy after the End of Nature*, MIT Press, Cambridge MA.
Wagner, Gernot. 2021. *Geoengineering: The Gamble*, Polity Press, Cambridge.
Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Krayer von Krauss, M. P. 2003. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17. https://doi.org/10.1076/iaij.4.1.5. 16466
Weinberg, Alvin. 1966. "Can technology replace social engineering?" *Bulletin of the Atomic Scientists*, 22, 10, pp.4–8. https://doi.org/10.1080/00963402.1966.11454993.

Whyte, Kyle Powys. 2011. The recognition dimensions of environmental justice in Indian country. *Environmental Justice, 4*(4), 199–205. https://doi.org/10.1089/env.2011.0036
Whyte, Kyle Powys. 2018. "Indigeneity in Geoengineering Discourses: Some Considerations." *Ethics, Policy & Environment* 21 (3): 289–307. https://doi.org/10.1080/21550085.2018.1562529.
Whyte, Kyle Powys, & Buck, Holly. 2021. Geoengineering and indigenous climate justice. In J. P. Sapinski, H. J. Buck, & A. Malm (Eds.), *Has it come to this? The promises and perils of geoengineering on the brink* (pp. 69–81). Nature, Society, and Culture. 2020 Rutgers University Press.
Wieners, Claudia E, Ben P Hofbauer, Iris E De Vries, Matthias Honegger, Daniele Visioni, Hermann W J Russchenberg, and Tyler Felgenhauer. 2023. "Solar Radiation Modification Is Risky, but so Is Rejecting It: A Call for Balanced Research." *Oxford Open Climate Change* 3 (1): kgad002. https://doi.org/10.1093/oxfclm/kgad002.
Winkelmann, Ricarda, Jonathan F. Donges, E. Keith Smith, Manjana Milkoreit, Christina Eder, Jobst Heitzig, Alexia Katsanidou, Marc Wiedermann, Nico Wunderling, and Timothy M. Lenton. "Social Tipping Processes towards Climate Action: A Conceptual Framework." *Ecological Economics* 192 (February 2022): 107242. https://doi.org/10.1016/j.ecolecon.2021.107242.
Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109 (1): 121–36.
Winsberg, Eric. "A Modest Defense of Geoengineering Research: A Case Study in the Cost of Learning." *Philosophy & Technology* 34, no. 4 (2021): 1109–34.
Young, Iris Marion. *Responsibility for Justice*. Oxford Political Philosophy. Oxford ; New York: Oxford University Press, 2011.
Young, Mark Thomas. "Now You See It (Now You Don't): Users, Maintainers and the Invisibility of Infrastructure." (2021) In *Technology and the City*, edited by Michael Nagenborg, Taylor Stone, Margoth González Woge, and Pieter E. Vermaas, 36:101–19. Philosophy of Engineering and Technology. Cham: Springer International Publishing, 2021. https://doi.org/10.1007/978-3-030-52313-8_6.

About the Author

As of September 2023, Benjamin Hofbauer is a postdoctoral researcher at the Research Institute for Sustainability in Potsdam, Germany. He completed his PhD between September 2020 and February 2024 at Delft University of Technology, the Netherlands. Benjamin's work is influenced by ideal and non-ideal political theories of justice, ethics and metaethics, as well as various strands within the ethics and philosophy of technology scholarship.

Benjamin has a Bachelor's degree in translation studies and literature from the University of Graz (Austria), as well as Master's degree in political, economic and legal philosophy from the University of Graz (Austria) and the Ruhr University Bochum (Germany).

Acknowledgements

Before diving into the acknowledgments, I would like to express how incredibly lucky and privileged I feel for having been given the opportunity to work and be paid as a researcher full-time. I never want to take this trust and the resources put into me for granted. While any PhD process has its ups and downs, I never once wished to be in any other position, doing anything else than this.

At the same time, I understand that people experience PhD trajectories differently, and that many struggle for a variety of reasons. Accordingly, I am aware that my experience is inevitably shaped by the privilege that comes with being a white European male at a European institution, and I have witnessed how colleagues who are not granted that kind of privilege have faced structural and individual injustices. Science, and particularly reflective undertakings such as philosophy and ethics, are only as good as the values that underpin them. Institutions that prize themselves in producing knowledge therefore continuously need to be reflective of and work against all forms of structural injustices, actively incorporating decolonial ways of education, stopping sentiments of white supremacy and misogyny in their tracks, and engaging in non-heteronormative modes of thinking and acting. Importantly, the struggle for social justice and equity is not a one-off policy change, but rather a continuous fight of opposition, subversion, and emancipation.

The last 3,5 years were intense and uniquely meaningful. Having moved countries three times, from Austria to Belgium, Belgium to the Netherlands, and the Netherlands to Germany, there was quite some turbulence and little time for me to physically settle. While the constant movement was inspiring in its own right, it also created a sense of urgency that was, at times difficult to deal with – a sense of urgency insinuating that if you are not moving, you are falling behind. It is particularly in these moments of difficulty that I always found reassurance throughout my family, friends, and loved ones. I would like to thank all the beautiful people whom I could lean onto in times of distress, who would lend me their ear and an understanding heart when I needed to vent, who would simply *be there*. I would not be able to manage the many curious movements and unpredictabilities of life, let alone finish a dissertation without you. I would therefore like to thank my brothers, Augustin and Christian, and my mom for always providing me with a home to return to. I also thank my oldest and always reliable friends, Matthias, Thomas, Antonio, Michael, Markus, who always cheered for me and made use of their humor when I was struggling. I would also like to thank Helena for always being there, asking how my work is going. Finally, I want to thank my partner, Adrienne: We may have only met in the final stretch of my PhD, but this time turned out to be one of the most significant ones in my life. Your patience, care, and love have eased even the most anxiety-laden phases of my journey, and I am incredibly grateful to have shared them with you. I want to thank all of you for being in my life.

I further had the privilege of being supervised by Behnam Taebi, Udo Pesch, and Ibo van de Poel, who’s knowledge and indeed wisdom is only matched by their kindness and empathy. Not once have I felt unduly pressured or left out of sight, and whenever I needed a quick response, feedback, or any form of support, I was sure to receive it. Ibo, Udo, and Behnam, have been an inspiration for how I would like to further develop as a researcher and a person, wherein respect and understanding for your colleagues are the central tenets of all your engagements.

I would also like to emphatically thank all my colleagues from the ESDiT consortium, who have been such excellent sparring partners in debate, and helped me immensely throughout the writing and thinking process. Ultimately, philosophy and ethics are communal endeavours and I could not have hoped for a better community to be part of. Importantly, I would also like to express my gratitude for all the organizational and administrative support I have received throughout the last years. Having had the pleasure of working alongside some of the administrators and management

staff, I learned a great deal about timely organization, and most crucially, the delicate balance of getting people to do things on time without making them feel pressured.

Samenvatting

In dit proefschrift worden de ethische uitdagingen behandeld die worden opgeroepen door een mogelijk onderzoeksprogramma voor zonnestralingsbeheer door middel van de injectie van aerosolen in de stratosfeer (Stratospheric Aerosol Injection; SAI). Deze ethische uitdagingen betreffen niet alleen kentheoretische obstakels met betrekking tot het onderzoeksproces zelf, maar ook maatschappelijke vraagstukken inzake rechtvaardigheid en de waarde van de natuur. In de scriptie wordt stilgestaan bij een veelheid van instrumenten en benaderingen voor het beoordelen en mogelijk beheersen van de risico's en onzekerheid die door SAI-onderzoek worden opgeroepen en op de maatschappelijke implicaties ervan. De methodologische benadering berust hoofdzakelijk op ethische en filosofische analyses en overwegingen, terwijl de belangrijkste conclusies zijn vervat in discursieve argumenten en normatieve bespiegelingen.

SAI is een vorm van klimaatengineering waarmee wordt getracht de opwarming van de aarde terug te dringen door het reflectieniveau van de atmosfeer te verhogen met behulp van de injectie van weerkaatsende partikels (aerosolen) in de stratosfeer. Alleen al de mogelijkheid om onderzoek te doen naar een technologie waarmee actief zou worden ingegrepen in het globale klimaat, is uiterst omstreden en heeft geleid tot verhitte debatten onder experts. Het ontwerpen van een onderzoeksproces voor een zo polariserende technologie als SAI roept onvermijdelijk fundamentele morele kwesties op, waarbij vraagstukken rond globale rechtvaardigheid, democratie, de waarde van onze verhouding met de natuur en de maatschappelijke gevolgen van technologische innovatie met elkaar zijn verweven. Gezien deze vérstrekkende consequenties ga ik in deze scriptie van de aanname uit dat SAI een zeer ontwrichtend idee en een zeer ingrijpende technologie is die bestaande maatschappelijke waarden en instituties ter discussie zou kunnen stellen en kunnen ondermijnen. Op grond daarvan worden in de scriptie een reeks filosofische onderzoeks- en beoordelingsmethoden aangedragen om eventuele regulerende programma's met betrekking tot SAI-onderzoek te voorzien van de benodigde ethische overwegingen en kaders.

De scriptie is gestructureerd in vier hoofdstukken, waarin afzonderlijke maar onderling verbonden vraagstellingen worden behandeld. Hieronder volgt een overzicht van deze onderzoeksvragen, naast een korte beschrijving van de beantwoording ervan.

> Centrale vraagstelling: *Hoe zou de wenselijkheid van onderzoek naar SAI beoordeeld kunnen worden, gezien de potentieel uiterst ontwrichtende invloed ervan op instituties, het milieu en maatschappelijke normen en waarden?*

Om deze centrale vraagstelling te beantwoorden zijn vier deelvragen geformuleerd, die in de vier hoofdstukken van de scriptie worden beantwoord.

> Deelvraag 1: *Wat zijn de ethisch relevante onzekerheden die binnen de context van een SAI-onderzoekprogramma zouden kunnen optreden en hoe kan met deze onzekerheden rekening worden gehouden?*

In dit eerste hoofdstuk worden de kernprincipes van het onderzoek naar SAI en de daarmee gepaard gaande ethische kwesties verkend. Daarbij ligt de nadruk op de behandeling van een herziene notie van 'normatieve onzekerheid', d.w.z. de onzekerheid die optreedt binnen tweeslachtige, moreel pluralistische besluitvormingscontexten op basis van multi-agentsystemen. Ik voer aan dat de besluitvorming omtrent de ontwikkeling en mogelijke toepassing van SAI binnen de context van de klimaatverandering een schoolvoorbeeld is van normatieve onzekerheid. Participatieve rechtspraak, erkenning en zowel reflexieve als adaptieve bestuursvormen worden opgevoerd als mogelijke benaderingen om in ethische zin rekening te houden met normatieve onzekerheid.

> Deelvraag 2: *Welke risico's en ontwrichtingen zouden zich als gevolg van een gecoördineerd SAI-onderzoeksprogramma op systemisch niveau kunnen voordoen en hoe zouden deze in ethische zin beheerst moeten worden?*

In het tweede hoofdstuk verbind ik de regulering van SAI-onderzoek met het concept van systemische risico's. Een van de kernproblemen met betrekking tot SAI-onderzoek is dat het tot verminderde politieke, institutionele en maatschappelijke steun voor radicale mitigatiemaatregelen ter bestrijding van de klimaatverandering kan leiden, een risico dat door sommige experts wordt omschreven als 'mitigatie-afschrikking'. In dat opzicht kan SAI-onderzoek worden afgespiegeld als een systemische bedreiging van de energietransitie en de daartoe behorende infrastructuur. Anders dan traditionele risicoconcepten zijn systemische risico's potentieel nadelige gevolgen die binnen de complexe infrastructurele en systemische context van moderne samenlevingen met elkaar zijn verstrengeld. Het zijn 'trigger points' die kunnen leiden tot de ineenstorting van hele systemen, waarbij de crash van de hypotheekmarkten in 2008 en de coronaviruspandemie als twee recente voorbeelden genoemd kunnen worden. Door de risico's van SAI-onderzoek als systemische risico's te beschouwen kunnen we het kernprobleem van mitigatie-afschrikking beter begrijpen en de beoordeling en beheersing ervan verhelderen. Daarmee verkennen we ook twee mogelijke aanvullingen op het kader voor systemische risico's, in de vorm van moreel dynamisme en betekenisvolle participatie en erkenning. Met behulp van deze ethische concepten, die in de filosofie en ethiek van technologieonderzoek zijn ontwikkeld, wordt het kader voor systemische risico's holistischer en kan nader worden ingegaan op kwesties van rechtvaardigheid en waarde met betrekking tot SAI-onderzoek en natuurwetenschappelijk en technologisch onderzoek in het algemeen.

> Deelvraag 3: *Hoe kunnen we een inschatting maken van de gevolgen die een onderzoeksprogramma voor SAI mogelijk zal hebben voor maatschappelijke structuren en instituties?*

Uitgaande van de normatieve onzekerheid en systemische risico's die in hoofdstuk 1 en 2 zijn behandeld, wordt in het derde hoofdstuk ingegaan op de ethische relevantie van de maatschappelijke dynamiek met betrekking tot SAI-onderzoek, en dan met name op het fenomeen van maatschappelijke en technologische 'carbon lock-in'. Ik verbind dat concept met noties als controle, omkeerbaarheid en innovatieve ontwrichting, met het argument dat we op grond van benaderingen als het voorzorgsprincipe kunnen inschatten welk soort innovaties ethisch aanvaardbaar is. Maar het voorzorgsprincipe, in welke gedaante dan ook, zal gepaard gaan met aanzienlijke beperkingen in de strijd tegen de klimaatverandering en zal uiteindelijk geen antwoord kunnen geven op de fundamentele vraag welke maatschappelijke waarden door nieuwe en opkomende technologieën afgedwongen *zouden moeten* worden. Op grond daarvan kom ik tot de conclusie dat de mogelijkheid dat SAI-onderzoek tot enige vorm van maatschappelijke en institutionele carbon lock-in zou kunnen leiden, op zichzelf geen ethisch probleem is. De aandacht zou veeleer moeten uitgaan naar welk soort waarden in zo'n lock-in zouden zijn vervat en in hoeverre die waarden omkeerbaar zijn en zouden moeten zijn.

> Deelvraag 4: *Welke gevolgen zou een SAI-onderzoeksprogramma hebben voor complexe maatschappelijke waarden, zoals duurzaamheid?*

In het vierde en laatste hoofdstuk verken ik mogelijke scenario's of 'vignetten' van technologisch-morele verandering als gevolg van de toepassing van een SAI-onderzoeksprogramma. Om te beginnen beschrijf ik verschillende opvattingen over de waarde van duurzaamheid, waarna ik verken welke invloed SAI-onderzoek zou kunnen hebben op de conceptualisering en maatschappelijke relevantie van die waarde. Met deze benadering probeer ik de praktijken van technologisch-morele verandering en scenarioplanning met elkaar te verbinden. Daarbij wordt niet zozeer gepoogd om definitieve voorspellingen over waardeverandering en -ontwrichting te doen, maar veeleer om toekomstscenario's te verkennen, waardoor wetenschappers en ook het bredere publiek en beleidsmakers de speelruimte wordt geboden om normatieve kernvragen te stellen over het soort maatschappelijke verandering dat we zouden willen bevorderen en over de wijze waarop bestaande wereldbeelden door de ontwikkeling van bijvoorbeeld SAI ter discussie gesteld zouden kunnen worden.

De vier onderzoeksvragen werpen licht op de noodzaak voor dynamische en geïntegreerde ethische beoordelingskaders voor de regulering van het onderzoek naar SAI en andere onderzoeksgebieden. Het kernargument gaat uit van het idee dat de beoordeling van technologische ontwikkelingen synchroon moeten lopen met de onderzoeks- en ontwikkelingsfase van nieuwe en opkomende technologieën, en dat de wisselwerking tussen het innovatieproces en maatschappelijke waarden aan de grondslag ligt van elke ethische onderzoeksregulering. Daarnaast is het van groot belang inzicht te krijgen in en rekening te houden met de morele context van de risico's en onzekerheid die met dergelijke processen gepaard gaan, niet alleen om met voorzorg en bescheidenheid te kunnen handelen, maar ook om de mogelijkheden van innovatie voor het teweegbrengen van positieve verandering en ontwrichting binnen vastgeroeste waardesystemen te herkennen. Buiten het thema van de regulering van SAI-onderzoek is deze scriptie in meer algemene zin relevant voor de ethische en filosofische beoordeling van opkomende technologieën en innovatie, voor het raakvlak van milieu-ethiek en de ethiek van technologieën en voor de ethiek van risico’s en de dynamiek van sociaal-technologische systemen.