Originally published as:

Stelzer, H., Schuppert, F. (2016): How Much Risk Ought We to Take? Exploring the Possibilities of Risk-Sensitive Consequentialism in the Context of Climate Engineering. - *Environmental Values*, *25*, 1, p. 69-90.

DOI: http://doi.org/10.3197/096327115X14497392134928

HOW MUCH RISK OUGHT WE TO TAKE? EXPLORING THE POSSIBILITIES OF RISK-SENSITIVE CONSEQUENTIALISM IN THE CONTEXT OF CLIMATE ENGINEERING*

Harald Stelzer
Institut for Advanced Sustainability Studies/University of Graz
Email: harald.stelzer@uni-graz.at

Fabian Schuppert
Queen's University, Belfast
Email: f.schuppert@qub.ac.uk

ABSTRACT

When it comes to assessing the deontic status of acts and policies in the context of risk and uncertainty, moral theories are often at a loss. In this paper we hope to show that employing a multi-dimensional consequentialist framework provides ethical guidance for decision-making in complex situations. The paper starts by briefly rehearsing consequentialist responses to the issue of risk, as well as their shortcomings. We then go on to present our own proposal based on three dimensions: wellbeing, fairness and probability. In the last section we apply our approach to a comparison of different climate policy options, including stratospheric solar-radiation management.

KEYWORDS

Climate engineering, consequentialism, risk, fairness, climate policy.

INTRODUCTION

As more and more time elapses without radical mitigation efforts on a global level, discussions intensify on the ethical, technological and political viability of planetary-scale climate engineering (CE) deployment. Within these debates on CE, it is striking to observe that very little is said about the merits of employing a consequentialist framework for assessing the ethical and moral status of different CE proposals. In fact, it seems as if consequentialism either is only used within climate economics, that is in the form of rather standard cases of cost-benefit analysis and cost-effectiveness analysis, or is criticised for being ill-equipped to deal with situations in which the consequences of our actions are laden with risks and uncertainties (Norcross, 1990; Lenman, 2000). While it is true that factors such as risk and uncertainty present a major challenge for correctly assessing the deontic status of actions, it is highly questionable whether consequentialism as such should be dismissed on these grounds, especially since outside of practical philosophy consequentialist reasoning plays an important role.

* Prepublished version. Now published: Stelzer, Harald and Fabian Schuppert (2016): How much risk ought we to take? Exploring the possibilities of risk-sensitive consequentialism in the context of climate engineering. *Environmental Values* 25 (1), 69–90. doi: 10.3197/096327115X14497392134928 Submitted 26 August 2014, accepted 27 May 2015

In this paper we set out to explore the possibility of a risk-sensitive multi-dimensional consequentialism, which is able to provide ethical guidance for our decision-making in complex situations such as rapid climate change. While there exist certain irreducible issues when it comes to assessing the deontic status of acts and policies in the context of uncertainty, we hope to show that employing a multi-dimensional consequentialist framework offers very plausible and situation-sensitive answers.

The paper starts by briefly rehearsing consequentialist responses to the issue of risk. We focus on so-called 'expected wellbeing approaches' and Martin Peterson's (2012; 2013) argument for a multi-dimensional consequentialist framework, which seems particularly suitable for dealing with complex decision-making situations and cases of risk. As we will argue, while traditional expected wellbeing consequentialism is too mono-dimensional, Peterson's account – despite its initial appeal – has some problematic features of its own. We then go on to present our own proposal which argues that consequentialists should: i) be value-pluralist and multi-dimensional in their axiology; ii) avoid probability-domination; iii) define guardrails and thresholds for permissible outcomes; and iv) focus on the overall satisfaction of a range of relevant thresholds (including probability) within a satisficing consequentialist framework. In the last section of the paper, we will apply our theory to a hypothetical decision on (not) employing stratospheric solar-radiation management (S-SRM), and briefly sketch the prospects and limits of our consequentialist framework.

I. CONSEQUENTIALISM AND RISK

In its simplest form consequentialism can be characterised as the doctrine that the deontic status of an act (including omissions, sets of acts and courses of action) depends only on consequences. While there exist many different forms of consequentialism with regard to the class of actions which fall under the doctrine and the range of right- and wrong-making properties, the key aspect for consequentialists dealing with risk and uncertainty lies in defining the kind of consequences on the basis of which the deontic status of an act is judged.[1] Traditionally, most consequentialists subscribe to one of two camps: either they are actualists or they are probabilists.

Actualists hold that only actual consequences matter. Whether or not an act is wrong depends solely on the actual consequences of that act (and often on the consequences that would have materialised if a different action had been taken). According to this view, hard-nosed actualist act-consequentialists might hold that pointing a half-loaded gun to the head of a friend and pulling the trigger is not morally wrong as long as no bullet is discharged and the friend is not harmed. There are of course many other reasons for which such an action might be considered blameworthy (see e.g. Thomson, 1986; Oberdiek, 2012) but speaking strictly in terms of actual consequences, no moral wrong was committed.[2] What matters for our discussion here is that actualists avoid the issues of risk imposition and risk-taking by focusing exclusively on actual outcomes or consequences. In cases of long-term risks, then, actualists can only assess the deontic status of a certain act, or policy, *ex post* in the distant future. In many cases, the fact that actualists are only able to pass judgment ex post is of little concern, since the relevant consequences materialise directly after, or very shortly after, the act. In the case of climate engineering, though, it might take decades before we know whether an intervention like S-SRM actually works, or whether it wreaks havoc. This raises the obvious issue of determining

[1] For an overview of consequentialism and its different battlegrounds, see Sinnott-Armstrong (2012), Darwall (2003), and Smart and Williams (1973).
[2] Not all actualist consequentialists would agree with this assessment, since actualist rule-consequentialists might argue that general careless (or risky) behaviour does actually lead to overall negative consequences.

at which point in time *t* we could, should or ought to assess the relevant consequences for defining the deontic status of an act, an issue which is not only of importance for actualists.

Probabilists, meanwhile, hold that probable consequences matter. Probabilists come in many shapes and forms, but the most relevant for our discussion here are those which hold that (subjectively or objectively) expected consequences matter.[3] The standard case for this variant of probabilism is expected wellbeing maximisation (see e.g. Smart, 1973; Jackson, 1991).[4] Probabilist consequentialists of the expected wellbeing variety carefully weigh the possible outcomes of an act, that is, an act's best and worst possible outcomes, and factor in the probability for each outcome. Thus, expected wellbeing consequentialists would regularly choose an act which produces relatively good outcomes with a high probability over acts which either produce very good or very bad outcomes with a low probability for the very good outcome. Probabilist consequentialists are thus able to easily differentiate between risky acts with potentially catastrophic outcomes and relatively safe bets.

As Peterson (2013: 104 & 107) points out, however, probabilist expected-wellbeing consequentialists tend to discount optimal outcomes due to their mono-dimensional value ordering which is virtually probability-dominated, since they focus only on the most likely outcome(s). That is to say, for expected wellbeing consequentialists it turns out that in many cases it is first and foremost probability which influences the deontic status of a given act. While to be mono-dimensionally probability-dominated might be considered unobjectionable in cases in which probabilities are with virtual certainty objectively-defined, being probability-dominated seems rather problematic in cases in which probabilities are less reliable. For expected wellbeing consequentialists, there does not seem to exist a difference in weight between probabilities based on standard procedures (e.g. the flip of a coin) and multi-variable calculations with significant error margins (e.g. predictions for the effects of large-scale geoengineering efforts). Part of this problem seems to stem from taking expected wellbeing and probabilities as compatible features of a mono-dimensional ordering system.

In response to these issues, Peterson (2012; 2013) presents his account of a multi-dimensional consequentialism. He argues that his account offers advantages over both mono-dimensional actualism and wellbeing consequentialism because of the introduction of three separate moral dimensions: individual wellbeing, equality and risk. Peterson (2013: 105–9) holds that these three dimensions are irreducible and incommensurable, which means that they cannot be collapsed into a single scale. Furthermore, he (2013: ch. 2) argues for accepting a distinctly non-binary conception of moral rightness. In other words, acts are not either right or wrong, but their rightness or wrongness comes in degrees.[5] While we very much agree with the basic tenet of Peterson's multi-dimensional consequentialism, that is, its commitment to three irreducible moral dimensions, and while we agree at least in part with the ideas of incomparability and degrees of moral rightness, on closer inspection we disagree with Peterson's account on two important scores. First, Peterson's actualist theory values optimal outcomes above everything else, which means that it is maximisation-dominated; and second, Peterson's theory gives us a particular understanding of incomparable sets of degrees of rightness, which in conjunction with the adoption of weighted randomisation as a decision-making procedure, renders Peterson's theory less informative and action-guiding than it could be. While our main

[3] Whether one specifies expected consequences on the basis of subjective or objective probabilities can obviously make a large difference in practice. For our discussion here, however, this is a point we can leave aside for now.
[4] In this paper we will ultimately defend a satisficing account of consequentialism. For our discussion here, however, we go with the standard probabilist view, i.e. expected wellbeing maximisation.
[5] An exception to this rule might be when we have an option on the table which is pareto-optimal with respect to all dimensions (i.e. the option scores best in comparison with the other available options in all three dimensions). In this case we may call this option simply right.

disagreement with Peterson's multi-dimensional consequentialism is with regard to the first point (i.e. maximisation-domination), the disagreement over the latter point is also of significance, as our hypothetical case-study will hopefully show. Let us briefly explain both points.

First, Peterson's theory offers implausible answers in cases like Jackson's (1991: 462–3) example of the treatment dilemma:[6]

A doctor must decide on the correct treatment for a patient who has a serious skin complaint. Careful consideration of the literature has led her to the following options. B will relieve the condition but will not completely cure it. One of A and C will completely cure the skin condition; the other will kill the patient, but there is no way she can tell which of the two is the perfect cure and which is the killer.

While Jackson argues for option B on probabilist grounds, according to Peterson, as consequentialists we should preserve the actualist's focus on best possible outcomes, which means that in the case described above treating the patient with B must always be considered wrong, since no matter whether A or C is the right cure, only A or C can deliver the optimal outcome. Hence Peterson argues that A and C should both be considered right to a degree and wrong to a degree, while B should be simply dismissed. This assessment, however, seems both counterintuitive and misguided, since it would ask policymakers to play Russian roulette rather than to choose a reasonably safe (even though not perfect) option. A moral philosopher might have the luxury of assessing the situation ex post and proclaim that killing the patient was wrong, but that does not help the patient, who was killed rather than relieved of a medical condition in lack of a better and *safer* cure. Peterson's multi-dimensional consequentialism therefore seems simply unable to deal with scenarios that are both haunted by uncertainty and very risky. As we will argue below, adopting a multi-dimensional expected wellbeing approach in conjunction with threshold levels for establishing moral permissibility could avoid the counterintuitive outcome discussed above.

Second, even if we leave extreme cases like the treatment dilemma aside, and focus on situations in which we simply have a range of options which score differently with respect to wellbeing, equality and risk, Peterson's account might leave us unsatisfied. As we said at the beginning of this paper, we want to explore the possibility of presenting a risk-sensitive consequentialism which can also provide at least some ethical guidance for decision-makers. This is not to say that our account will be able, or should be able, to say for every situation *S* what the morally right course of action is. Instead we want our theory to define a range of theoretically (im)permissible options and a matrix which allows decision-makers to see how different options can produce different outcomes across three irreducible moral dimensions. This is exactly where Peterson's theory partially struggles. Peterson is committed to both a particular understanding of the irreducibility and incomparability of the three moral dimensions, and a particular conception of degrees of moral rightness, which means in most complex cases we will end up with statements of the following kind:

Option A is 20% right with regard to wellbeing, 30% right with regard to equality and 50% right with regard to risk. Option B, meanwhile, is 30% right with regard to wellbeing, 50% right with regard to equality and 20% right with regard to risk, while option C is 50% right with regard to wellbeing, 20% right with regard to equality and 30% right with regard to risk. Peterson suggests that, taking these degrees of rightness into account, we should adopt a weighted lottery as the appropriate decision-making tool.

The problem with statements like this and the proposed decision-making procedure is that they give very little *ethical* guidance to decision-makers. In fact, if we followed Peterson's proposal, no *real*

[6] Interestingly enough, Peterson (2013: 106–7) uses this example to defend his view. We, however, think his interpretation is misguided, due to his theory's maximisation domination.

decision would need to be made, as the actual decision is the outcome of a weighted lottery. Once we have agreed on the values in the matrix, we just perform a weighted lottery and go with whatever the result is. While we should of course consider each and every option which might be the right one (which is exactly what the weighted lottery is designed to do), this does not exempt us from the responsibility to argue for or against a certain option, including argument on the basis of cross-dimensional considerations. Even though the different moral dimensions do not collapse into a single one and are strictly speaking non-comparable, we can have weighty reasons to prefer – for example – wellbeing to equality, or to take higher risks in certain situations. The incomparability of the dimensions (in a strict sense) does not necessarily lead to a position that prohibits trade-offs between different dimensions based on moral and other reasons. Peterson's account seems to exclude careful deliberation over the different options on the table, and other reasons we might have for preferring A over B. Instead, all possible considerations are expected to have been in the matrix, which means Peterson's account assumes that the incomparability of the three separate moral dimensions is absolute, not only with regard to moral rightness (a claim with which we agree) but also with regard to the decision-making: that is, it prevents us from using other reasons for making a decision or for taking a particular option of the table (for instance, because it violates an absolute threshold).

To be clear, Peterson does not claim that any multi-dimensional consequentialism must subscribe to either absolute incomparability or weighted randomisation as a decision-making procedure. Our disagreement with Peterson's theory is therefore internal, and with regard to negotiable features of his multi-dimensional consequentialism:[7] that is, with regard to how one should properly conceptualise a multi-dimensional consequentialism, not whether such an account is tenable at all. As Peterson (2013) himself points out, embracing his particular understanding of moral degrees and his proposed decision-making procedure is in a sense optional and independent of the deeper commitment to multi-dimensional consequentialism as such.[8] While Peterson's multi-dimensional consequentialism is certainly risk-sensitive, we want to argue that adopting a satisficing expected wellbeing account allows us to establish clearer boundaries of moral permissibility, that it avoids the problems associated with Peterson's actualist account, and that it lends itself to adopting different decision-making procedures.

II. SUFFICIENCY-RESTRAINED MULTI-DIMENSIONAL CONSEQUENTIALISM

From what we have said so far, we can already gather a long list of requirements which we want our alternative conception of consequentialism to fulfil. First, it should be risk-sensitive. Second, it should neither be dominated by probability, nor by maximisation. Third, our consequentialist framework should provide ethical guidance for practical decision-making. Admittedly, this is a rather tall order.

Based on our considerations in Part I, we want to stick to the idea of a multi-dimensional consequentialism. Like Peterson, we will define three basic moral dimensions: individual wellbeing, probability and fairness.

Wellbeing matters because it tells us how people get affected, and which gains or losses an outcome will bring about. Just like Peterson (2013: 50), we take the dimension of wellbeing to be concerned with every individual's wellbeing, so as to respect the separateness of persons. This is an important

[7] By calling these features 'negotiable' we want to highlight the fact that the most radical and important aspect of Peterson's multi-dimensional consequentialism, namely, its commitment to three irreducible and morally incomparable moral dimensions remains in our proposal intact.
[8] We are grateful to an anonymous reviewer for encouraging us to clarify this point.

point, since most standard consequentialist views see individuals only as substitutable containers of wellbeing, which is actually one of the reasons why critics of satisficing consequentialism, such as Tim Mulgan (2001), assume that satisficers can get away with murder, a conclusion that is simply wrong if one's wellbeing dimension respects the separateness of persons.[9]

In terms of wellbeing we focus on whether an option enables or disables the realisation of a person's fundamental interests.[10] The fulfilment of those interests establishes a threshold of wellbeing. Framing wellbeing in this way means that one should consider options which would lead to people falling below such a threshold (so that they are unable to satisfy their fundamental interests) morally impermissible, as far as such consequences are predictable and could easily be avoided by some other course of action. With this interest-based approach we distance ourselves from the problematic utility assumption underlying many consequentialist positions in economic theory, as well as their attribution of monetary values to all benefits and burdens, including human life, physical security, subsistence and health.

The protection of people's fundamental interests might be considered a rather uncontroversial minimum standard. Whether such a threshold is demanding enough, considering that in complex situations (in which no easily available pareto-optimal option exists) mere basic-interest-fulfilment would be considered good enough (from the viewpoint of moral permissibility), is a question which certainly might cause some controversy. Moreover, in many complex decision scenarios it is unlikely that any option will be able to guarantee that no person's basic rights are violated, hence it is important to allow for some leeway. Furthermore, in most realistic scenarios, we will only have a rough idea about what kind of consequences might materialise if we choose a particular option. Due to issues of complexity and uncertainty, we will only be able to sketch likely outcome scenarios. These scenarios can be described as consequence-bundles understood as groups of possible consequences according to certain features (i.e. catastrophic ones and fairly positive ones). Depending on the uncertainty and complexity of the choice situation, these consequence-bundles can be more or less spread-out and wide. Positive wellbeing gains in such bundles could be construed as chances, while wellbeing losses could be understood as risks. What we have to do, then, is to use our best available predictive tools in order to sketch what chances and risks a certain option might offer.

Probability matters because it tells us how good or bad our chances are that a certain consequence, or in most cases bundle of consequences, will materialise, and how large the uncertainties that we are facing are. Take the case of having in a situation *S* two options, O_1 and O_2. O_1 brings about consequence-bundle A or bundle B, where A meets our wellbeing threshold, while B fails to satisfy it. O_2, meanwhile, will either result in consequence-bundle C or bundle D, where C vastly exceeds our wellbeing threshold (thus promising great wellbeing gains), while D fails to satisfy it. If we were not to take probabilities into account, O_2 would be clearly more attractive than O_1. Once one takes probabilities into account, however, this might well change, for instance if the probability distribution for O_1 was (0.5 | 0.5) and for O_2 (0.02 | 0.98). Now the potential wellbeing gains of choosing option O_2 seem vastly less attractive, since the probability of bundle C coming about is only 0.02.

It is important to be clear about what the probability dimension refers to: technological feasibility, political and economic feasibility, or the probability of attaining a particular outcome or consequence-

[9] We will return to the issue of satisficing, and Mulgan's criticism thereof, below.
[10] Fundamental interests refer to the moral interests people have for leading a decent life, or good life (depending on the account of fundamental interests to which one subscribes). Fundamental interests ground basic rights in the things and goods people need for leading a decent life or good life. For an account of fundamental interests, see Schuppert (2013).

bundle within a chosen option. In our view, the probability dimension can only be concerned with the latter and not with the former. Which options $O_1 \ldots O_n$ are at all on the table, which is an issue of political and economic feasibility, is another question. In fact, this is an issue which has to be dealt with prior to deciding which option one ought to choose, since the description of each and every option (and indirectly thus of every consequence bundle) is based on a range of political and economic feasibility assumptions. Once we know which options are on the table, we need to compare the available options with regard to their performance in all three dimensions: wellbeing, fairness and probability. Our multidimensional framework therefore provides a second-order evaluation of options based on normative criteria. Feasibility constraints, however, come into play again, once we have done our normative assessment of all available options; after all, it might well be the case that the best option (normatively speaking) proves utterly unfeasible because of political and economic circumstances.

Our basic starting point for factoring risk into our framework is therefore the assessment of every considered option's possible consequence-bundles, and how they perform with regard to the relevant thresholds. In contrast to the example given above, however, in many cases we will not be able to assign numerical values to a bundle's probability and consider the full range of possible outcomes of an option; moreover, we will have to carefully defend the probabilities we assign to particular consequence-bundles, and make sure that we reflect existing uncertainties within our probability dimension.

Furthermore, as we saw in our (and Peterson's) discussion of expected wellbeing theories, if one were to apply a mono-dimensional scale in conjunction with a focus on probable (and also possible) consequences, one's decision-making would become easily dominated by probabilities: that is, one would tend to focus only on the most likely outcome. One way to avoid the problem of probability domination is to use a multi-dimensional framework, which does not allow for collapsing probability and wellbeing into one ordering scale. Employing such a multi-dimensional framework within a satisficing theory has the further advantage that one determines separate probability thresholds, since for positive consequence-bundles we want probability to be high (i.e. the outcome to be likely), while for negative consequence-bundles we want probability to be low (i.e. the outcome to be unlikely). Establishing different thresholds allows us to exclude those options which feature consequence-bundles that score very well in terms of wellbeing, but which come with very low and unreliable probabilities for these positive bundles. The same goes for the other way round, which means that spelling out probability and uncertainty while establishing thresholds also allows for identifying alarmist doomsday scenario outcomes, which are so extremely unlikely that they should not over-proportionally influence our decision-making.[11]

Even though our approach is therefore not probability dominated, it is still risk sensitive, as it takes the risks of different options into account by including the risks that we ascribe to different options in our wellbeing function, and by mapping uncertainty in our probability function. By setting a certain threshold for wellbeing as well as for the probabilities of all these outcomes, such an approach can be tuned to be more or less risk-sensitive, and it can be done so with respect to the underlying problem situation, and what is at stake. One of our core assumptions is therefore that the moral (im)permissibility of an option may change when we consider it in different scenarios. Based on these

[11] That is not to say that satisficing consequentialists like us must necessarily treat all doomsday scenarios as cases of unrealistic alarmism. Instead it allows us to put things into perspective, since keeping probability as a separate indicator is less likely to obscure our ordering than a mono-dimensional scale. Basically, setting a risk threshold prevents us from extreme risk-aversion and extreme risk-taking.

considerations, we assume for our multi-dimensional satisficing consequentialism: a) that we need to have probability thresholds that neither arbitrarily ignore unlikely or unpredictable catastrophic outcomes, nor are dominated by very unlikely scenarios;[12] and b) that such thresholds themselves can be variable, depending on the scenario we are dealing with and the possible alternatives available.

What matters beside the consequences and their probability is also how the costs and benefits of an option are distributed. We therefore include a third moral dimension: fairness. As Peterson points out (2013: 74), fairness is simply broader than equality, since it takes considerations beyond pure equality into consideration.[13] Fairness matters, as it tells us whether people get equally affected, or whether some benefit at the expense of others. For the dimension of fairness, we operate with the assumption that an outcome (including distributions) can only be considered fair if it does not arbitrarily disadvantage certain people, groups or generations: that is, if the distribution of benefits and burdens associated with an outcome track the relevant reasons. This means that according to our criterion of fairness, an act which intentionally produces wellbeing for one group at the expense of another group should be, in the absence of other (possibly overweighing) reasons, considered unfair.

We want in this paper to spell out the idea of fairness along the lines of a gradual scale, from 'very fair' to 'very unfair' (with a range of intermediate labels such as 'quite fair', 'somewhat fair', 'somewhat unfair', 'quite unfair'), assuming that 'somewhat fair' might be an acceptable threshold for most cases. In fact, based on the assumption that it is morally impermissible to gratuitously compromise the rights and interests of people regardless of their spatial and temporal placement, we suggest that the principles of wellbeing and fairness are to be applied both globally and intergenerationally. One significant advantage of this reading of wellbeing and fairness is that it safeguards the fundamental interests of the most vulnerable, i.e. future generations and the poor. As pointed out earlier, however, under the pressure of non-ideal circumstances a limited range of available options might lead in certain situations to a state of affairs in which we might have to settle for a somewhat unfair option, simply because no better options are at hand. This is a problem we will discuss during the application of our consequentialist framework to the issue of climate engineering.

What we advocate in this paper, then, is settling for a three-dimensional satisficing consequentialism, which sets separate threshold levels for each moral dimension. As satisficing consequentialism is controversial, let us make clear what kind of satisficing consequentialism we advocate, and why we take it to avoid the pitfalls commonly associated with satisficing.

First of all, it is important to note that we argue for satisficing consequentialism as a suitable strategy in complex decision-making situations with significant variations in both possible outcomes and assumed probabilities, and in which no pareto-optimal solution (across all three irreducible moral dimensions) exists. If there exists a clearly discernible pareto-optimal option across all three moral dimensions, this option *ought to be chosen*, since it is the only morally right one. If, however, no option exists that is pareto-optimal across *all three* dimensions, due to the dimensions' irreducibility

[12] In order to avoid arbitrariness, we obviously have to choose our probability thresholds on the basis of good, defensible reasons.

[13] According to Peterson, both equality and fairness are suitable categories to be used. As Peterson (2013: 74) puts it, 'multi-dimensional consequentialist[s] ha[ve] some freedom of choice' with regard to this question. We choose fairness because equal outcomes and equal procedures are not always fair, implied by virtue of the fact that factors other than equality can inform what counts as fair in particular circumstances.

and incomparability, no option can be labelled as being fully morally right. Instead several options might be morally permissible.[14]

Second, in order to establish which options *ought to be* considered morally permissible, we advocate satisficing in the form of threshold-level-setting for each moral dimension. Hence only options which satisfy the permissibility thresholds set for each dimension are to be considered morally permissible. Using the idea of satisficing as a way of establishing moral permissibility thresholds within a multi-dimensional consequentialist framework is thus entirely different from using satisficing as a general principle within a mono-dimensional scheme of what has been labelled 'blatant moral satisficing'(Slote, 1984; Mulgan, 1993). Blatant moral satisficing is simply too crude, since it allows for a range of counterintuitive actions, such as making people gratuitously worse off as long as they stay above the threshold (Bradley, 2006).[15] Our version of satisficing avoids these issues by i) valuing the separateness of persons, ii) disaggregating the consequences of an act along three irreducible moral dimensions, and iii) establishing a non-suboptimality condition which holds that acts which are in comparison to another available option pareto-suboptimal (i.e. worse across all three dimensions) are morally impermissible. The reason for the non-suboptimality condition is simply that if we were to choose a suboptimal option we would fail to bring about the best moral outcome from two comparable options, which clearly runs counter to the underlying rationale of our expected outcome consequentialism, since the idea of satisficing for establishing moral permissibility only comes into play once we have no comparable options available. Thus, in the following table, only options A and C are morally permissible even though B, too, satisfies all the established threshold levels. Option C, however, is across all three dimensions better than option B, which means that B fails to satisfy the non-suboptimality condition.

Table One

	Wellbeing	Probability	Fairness	Permissibility
Threshold value for moral permissibility	5	2	2	
Option A	8	3	3	morally permissible
Option B	6	3	3	morally impermissible
Option C	7	4	4	morally permissible

Third, if there is no option which meets all the threshold levels across the three dimensions, we face a situation of a problematic nature. We therefore need to carefully assess the potential negative consequences for wellbeing involved in each option, assess the probabilities, and compare the expected fairness and wellbeing outcomes so as to decide which option is morally permissible to choose in these non-ideal circumstances. That is, we assume here that even in dilemma situations one or more options are morally permissible. A different way of interpreting dilemma situations would be

[14] What will happen in cases in which no option satisfies the thresholds of all three dimensions is a question we will address below.
[15] Rogers (2010) presents a mono-dimensional form of satisficing consequentialism which avoids this problem.

to say that all options are morally wrong, but some are less bad than others. We find this second reading implausible. Naturally, in cases in which each option features a worst outcome which threatens to fail to satisfy all our sufficiency constraints, and if the probabilities in these cases are rather unreliable, it may be difficult to establish which option(s) should be considered morally permissible. We will come back to this problem in our example below.

What interests us in the idea of satisficing is therefore not that it is morally permissible for an agent to simply identify a 'good enough' option and to go for it. Rather, satisficing consequentialism is interesting as it includes the idea of some aspiration- or sufficiency-level, which can be understood as a threshold-level of moral permissibility which could be reached by different options (Pettit, 1984: 166).[16] By combining the idea of satisficing with a multi-dimensional approach, which assumes value-incommensurability and -incomparability, satisficing consequentialism provides us with a great tool for carefully unpacking the complex consequences of different options. The exact decision-making procedure for choosing an option is not fixed, though it seems sensible to assume that the expected overshoot (i.e. the wellbeing- and fairness-gains above the threshold), as well as the overall probability, should influence the decision. Instead of using a weighted lottery and leaving the decision up to chance, however, we propose that one of the morally permissible options is chosen on the basis of careful deliberation and possibly the consideration of other noteworthy reasons, external to the three moral dimensions discussed above. Moreover, because our account of multi-dimensional consequentialism focuses on expected outcomes which satisfy a set of particular moral permissibility thresholds, it avoids Peterson's issue of maximisation domination.

In a nutshell, our approach therefore leads to the following decision rules:
(1) If there is one option which not only satisfies all thresholds but which is also pareto-optimal across all three dimensions, this option *ought to be chosen*;
(2) If no pareto-optimal option exists, all the options which meet all the thresholds across all three dimensions are *morally permissible* as long as they satisfy the non-suboptimality condition;
(3) If no option satisfies all thresholds we face a dilemma, since no option satisfies the demands of our framework.
As mentioned above, however, even in dilemma situations we hold it to be possible to assess options so as to find out which options can be considered morally permissible in these non-ideal circumstances.

III. CONSEQUENTIALIST DECISION-MAKING AND CLIMATE ENGINEERING

As put forward in the Introduction, the development of our multidimensional consequentialism has been inspired by problems of climate change and the possible use of climate engineering techniques. One technique that seems of special interest to philosophers, due to the difficult problems it raises for decision-making under uncertainty, is stratospheric solar-radiation management (S-SRM), the albedo modification technique currently most discussed. S-SRM involves increasing the amount of aerosol particles in the lower stratosphere, as a means of increasing the reflection of sunlight beyond what is reflected by the naturally-occurring stratospheric aerosol layer. Particles could either be injected directly or formed via injection of precursor gases like SO_2, which are then converted into particles (Schäfer et al., 2015). Due to its different features, as discussed below, S-SRM makes an interesting case for applying our account as developed in Part II. This application, however, is to be seen as a

[16] Where such a threshold should lie is of course open to debate and the attractiveness of our framework somewhat depends on choosing appropriate threshold levels.

case study rather than as targeted policy advice. Our main concern here is to bring forward some merits and problems of our approach, also indicating the need for its further development.

For such an analysis, it is first important to decide on the alternatives, the options or scenarios to which S-SRM will be compared. For our analysis here we chose different forms of mitigation as well as unmitigated climate change, which we expect to lead to very bad outcomes, and the burdens of which will be generally heavier on those worse-off in terms of income, education or social status (Louis and Hess, 2008; Schneider et al., 2007).[17] Second, doing so, what time horizon we take into account seems relevant, as well as how we design the scenario. Without doubt the framing of the scenario will influence the attractiveness of the different options. For the following example, we will assume a time-frame of 2016 until 2100.

We will use the criteria we defined above: human wellbeing, probability and fairness. We will set the wellbeing threshold at the protection of people's fundamental interests. For answering the question if this threshold can be fulfilled by an option, we will not be able to describe the expected consequences in detail, but rather give a rough estimate of the expected outcomes. The probabilities indicate the chance of attaining a particular outcome or consequence-bundle within a chosen option. We will not try to assign numbers to the different options but rather work with probability ranges. Fairness, which plays an important role due to the spatially and temporally heterogeneous distribution of benefits, costs and risks, can range from 'very fair' to 'very unfair' (with a range of intermediate labels such as 'quite fair', 'somewhat fair', 'somewhat unfair', 'quite unfair'). The dimension of fairness takes into account the distribution of benefits and burdens on present and future people (more or less directly) affected by the options, which means that we assume a time-frame of roughly one hundred years. One important reason for choosing this time-frame is that the further we go into the future, the less we know about the state of the world, meaning that while the future effects of our policies or choices are always somewhat uncertain once we go too far into the future, we simply seem more or less clueless as to how things will look, and how strongly our present choices will bear on future lives and opportunities.

In the following table we sum up the options that we will compare: aggressive mitigation (AM), moderate mitigation (MM), unmitigated climate change (UCC)[18] and stratospheric solar-radiation management (S-SRM). It is important to note that by considering these four options we have already made two important decisions: namely, we have decided to include AM even though it is politically not very likely to happen, and we have included S-SRM despite concerns about its technological and political feasibility. The reason we have done so is simple: while we hold issues of political and technological feasibility to be important, in order to show the potential (and also the limitations) of our consequentialist multi-dimensional framework, we want to consider both AM and S-SRM. In the rest of the paper we will spell out this comparison in more detail and arrive at some preliminary conclusions.

Table Two

	Expected Wellbeing (until 2100)	Probability	Fairness

[17] For reasons of simplicity we exclude adaptation-options, even though for some they seem to be the most feasible, due to the low political feasibility of aggressive mitigation and the severe risks and uncertainties of S-SRM.
[18] We will not consider UCC in more detail below, but rather take it as a control case for the other options.

Aggressive Mitigation (AM)	negative effects due to still rising temperatures, but long-term basic-interests protection attainable; low risks for wellbeing due to climate-change-induced impacts	high probability of reaching expected outcome	quite fair
Moderate Mitigation (MM)	negative effects due to temperature rise above two degrees Celsius; increase of severe weather events; high risks due to climate-change-induced impacts	relatively high probability of reaching expected outcome, depending however on climate sensitivity as well as vulnerability and resilience of environmental and societal systems	quite unfair
Unmitigated Climate Change (UCC)	very negative effects; high temperature increase; severe effects globally	high probability of reaching expected outcome; chance of unforeseeable catastrophic outcomes	very unfair
Stratospheric solar radiation management (S-SRM)	(if it works) lessening of harms by climate change; while climate change risks would decrease, side-effects of deployment could be severe;	no clear probability of reaching expected outcome, due to scientific and societal uncertainties; chance of unforeseeable catastrophic outcomes	somewhat unfair

Aggressive mitigation (AM) – defined as a large-scale effective mitigation effort, starting now, and enabling us to stay within the two-degrees-Celsius target until 2100 – scores best by far in terms of wellbeing as well as fairness. Even though AM would not be able to guarantee the protection of all people's fundamental interests, as it would slow down temperature rise significantly, we expect, if deployed in time, this option would lead to outcomes that would most likely fulfil our wellbeing threshold to a higher degree than any other option available.

Still, it is not to be forgotten that AM would also come with costs, especially for those living in industrialised countries. How high these costs would be is open for debate. Nevertheless, it is clear that the burdens and benefits of AM would not be equally divided between all people on the planet, as based on the necessary technological transfer and the need of emission reductions, the costs will most likely rest upon the North, while the benefits will be on the South.[19] Still, this seems to be an almost fair distribution, especially if one factors in considerations of historical (in)justice, since the North has caused most greenhouse gas (GHG) emissions and has benefited most from past emissions, while the South is most vulnerable to climate-change-induced consequences and has (up to now) emitted only a small fraction of the existing GHG concentrations (Meyer and Roser, 2010). Furthermore, AM will also lead to the best outcome from the intergenerational perspective, as people today would bear the costs in order to benefit future generations, which will have to take most of the consequences of global warming (Meyer and Roser, 2009). It is also the fairest, as it goes against the root causes of the

[19] One should also consider that mitigation comes with many side-effects that might affect the most vulnerable, like the e-waste problem of the alternative energy sector, or justice questions raised by programs for the protection of tropical rainforests. It is important to deal with these problems in a fair and sustainable way.

problem. In short, AM is pareto-optimal and is therefore the right choice: that is, AM *ought to be chosen*.

When evaluating AM, however, we have to take one particular factor into account: it works best only within a certain time-frame. So our evaluation would change if we were to consider AM starting only in the year 2060, as by this point in time it may already be too late to avoid crossing certain temperature thresholds or climate tipping-points by mitigation alone. The chances of meeting our threshold of wellbeing are therefore bound to decrease the longer we wait with our decision.

Moderate mitigation (MM), defined as a slow decline of emissions by 2030, is politically speaking much more likely, but it is expected only to postpone some of the negative outcomes of climate change and/or make them only less bad. Even though mitigation efforts would still have to be substantial, they would most likely fall short of reaching the threshold of wellbeing. The probability of reaching the assigned consequence-bundle – based on our current understanding of climate models and integrated assessment models – would depend on climate sensitivity, as well as the vulnerability and resilience of environmental and social systems. The probability that the expected outcome materialises is therefore only relatively high. Even if we assume that future societies would be – due to resilience, technological progress and greater wealth – more adaptable to severe climate change, we expect lower overall levels of wellbeing in comparison with AM, especially from an intergenerational viewpoint. The spatial and temporal difference of the impacts of global warming and the 'backloadedness' make the distribution of costs and benefits in the case of MM quite unfair. MM thus fares much worse than AM, since AM scores better across all dimensions, even though in comparison to UCC, MM is more attractive across all dimensions.

For our evaluation of S-SRM, we use the same time-span as for the other options, even though we take it to be unreasonable to even start considering S-SRM before exhausting other options, or reaching a certain threshold, like a climate tipping-point or a certain temperature increase (e.g. two degrees Celsius). In order to non-arbitrarily assign values to S-SRM, we have to look more closely at its possible effects. Even if S-SRM works, that is, if it lowers the global mean temperature as expected, and by doing so counteracts some severe effects of global warming or even hinders the approach of a climate tipping-point, S-SRM is expected to have side-effects and potentially large risks: changes in local temperature, precipitation and climate patterns such as El Niño and monsoons; effects on regional food and water availability, potentially leading to droughts and famines; ozone depletion; changes in marine and terrestrial biological ecosystems. Furthermore, S-SRM alone will not reverse all of the adverse effects of rising emissions, like the acidifying effect of carbon dioxide on the oceans. Even if ongoing research may lead to ways of managing some of those risks and side-effects, temperature as well as precipitation differences between regions might increase with time, if S-SRM is not accompanied by large mitigation efforts. Furthermore, longer periods of S-SRM deployment would create a permanent threat due to a potential termination as the failure to maintain the aerosol counterforcing could result in abrupt and potentially very damaging warming (Ross and Matthews, 2009).

Due to the risks involved and the potentially diverse regional climate effects, S-SRM may harm or disadvantage some, while benefiting others (Ricke et al., 2010). The distribution of benefits and costs would not only depend on existing climate conditions but also on population density, economic development, and the vulnerability and resilience of ecological, economic and social systems. This could pose particular issues for global fairness, as S-SRM could transfer risks to the poorest countries. Those most vulnerable to climate change geographically and economically, often living at the

subsistence level, could also be those most likely to be negatively affected by uneven effects of S-SRM and have the lowest capacity to adapt to its consequences, despite being least responsible for global warming (Preston, 2012). S-SRM could also increase inequalities between generations if harms and risks are deferred to a later time (Goes et al., 2011; Burns, 2011). Whether the deployment of S-SRM would increase existing (or future) inequalities and historical injustices of climate change is, however, an open question, as it may also benefit some of the most vulnerable and poorest countries by reducing climate change-induced risks (Svoboda et al., 2011; Tuana, 2013).

Questions about the distribution of effects are even more important considering that S-SRM may fail, and in doing so worsen the harmful consequences of climate change rather than alleviating them. On the one hand, S-SRM could turn out to be of little value, once deployed and associated with severe negative side-effects which are not compensated by the effects of lowering the global mean temperature. If climate change is understood not just as a problem of mean global temperature rise, but more complex on account of its regional and local impacts, such a (non-catastrophic) failure of S-SRM seems possible (Robock et al., 2008). On the other hand, in particular non-linear internal feedback between various components of the climate system could result in bifurcations of the system and might therefore lead to abrupt shifts or transitions between states. It is possible that even smaller forcings in certain situations could lead to passing a critical threshold to an unstable condition (Tuana et al., 2012). This mere possibility of *unexpected catastrophic* consequences, which may be far worse than any of the expected ones, especially seems to undermine this option, and could jeopardise the ethical acceptability of S-SRM (Davies, 2010).[20]

We are aware that this outline of aspects linked to S-SRM opens up more questions than it provides answers. The normative evaluation of S-SRM is therefore complex and an open process. Based on the above-mentioned issues, however, we will give a preliminary assessment.

When assessing S-SRM we face one key problem: assigning probabilities towards S-SRM failure or success seems virtually impossible, based on the controversy in the scientific field and the severe uncertainties involved. We also have no idea about the likelihood of catastrophic outcomes, even though the mere chance of such S-SRM catastrophe clearly speaks against deployment. At the same time, we have to consider that in a scenario of unmitigated climate change, the complexity of the climate system and possible feedback processes might also hold some catastrophic 'climate surprises' (Bodansky, 2011). Even though a failure of deployment would very likely increase harms, in the case that S-SRM works and reduces other climate-change-induced risks, we hold it to be possible that our wellbeing threshold could be satisfied at least to some degree. For the moment this seems realistic, even taking into account the known potential side-effects. Benefits could also be increased if S-SRM is considered not as a single measure but in combination with other options, like increased mitigation efforts or greenhouse-gas-removal techniques. Doing so, one would be able not only to hinder and/or ease out some side effects, but also to decrease the amount and time of S-SRM deployment. Moreover, S-SRM might actually not be as unfair as moderate mitigation (MM) or unmitigated climate change (UCC), which do nothing or too little to support victims of harms by climate change, especially if it is coupled with compensations for those affected. How fair it is will depend, however, not solely on the success of the option but also on taking or transferring costs and risks. Even if people at the time of deployment would take at least some risks of S-SRM, the risks of long-term implementation and potential termination would be transferred to the future. Evaluating the fairness of

[20] For reasons of space we will ignore other potential negative consequences, due to the high conflict potential of such techniques.

S-SRM, it therefore seems justified to conclude that S-SRM will always lead to somewhat unfair outcomes.

At this point of our analysis, it would be a nice thing to be able to assign real numbers to all outcomes and to just sum up the scores of AM, MM, UCC and S-SRM in order to reach a conclusion. Two reasons, however, speak against such a simple aggregation. First, it would contradict our multidimensional account, and second, due to the speculative nature of our assessment and especially the uncertainty regarding the effectiveness of S-SRM, a straightforward comparison is not possible. What is clear, however, is that AM is better than MM, and MM is better than UCC. Moreover, due to the high probability we attach to AM generating outcomes which meet the thresholds and due to our more sceptical account of S-SRM, it seems prudent to choose AM over S-SRM. As our analysis also has shown, though, AM is an option which a) comes with an expiry date, and which b) might not even be on the table due to a lack of political will to engage in AM and to take the burden of its associated short-term costs. Thus, while our obvious recommendation is to engage in AM, the question arises of what to do in a case where AM alone is no longer an option for avoiding the considerable harms from climate change, and where additional adaption measures will not be able to hinder most of the climate-change-induced harms either.

To compare MM with S-SRM is much more difficult, precisely because our analysis of S-SRM is based on a range of assumptions built on less-than-reliable probabilities and guesses. Generally speaking, at this point in time, and based on our current knowledge, we consider it unreasonable to even start considering S-SRM before exhausting other options, including MM, or reaching a certain threshold, like a climate tipping-point or a certain temperature increase (e.g. two degrees Celsius). Thus, if we had to make a once-and-for-all choice now, our framework speaks in favour of MM over S-SRM. Since our assessment might change because of increased knowledge regarding S-SRM and its effects (and also because other options might have run their course), or a higher climate sensitivity, we do not think that a strict moratorium on S-SRM research can be defended on the basis of our framework, unless one wants to take S-SRM as an option completely off the table. This, however, seems not the advice to give considering UCC.

CONCLUSION

Applying our approach outlined in Section II to the analysis of climate change policy options shows some strengths as well as some weaknesses of our multi-dimensional consequentialism. Uncertainties remain a great challenge, assigning values to the expected consequences of an option can be difficult, and defining threshold-levels can cause controversy, too. One further problem we are well aware of is the possible confusion between our understanding of probability with questions of political and economic feasibility. Even though we focus during our normative evaluation on the probability of a certain consequence-bundle once an option is chosen, the question of which options are on the table and which are implemented in the end is at least partly determined by feasibility considerations.

Despite all the issues, we think that our consequentialist framework has great potential for dealing with complex decision-making scenarios, and that our analysis of Table Two has shown how our framework can generate guidance for decision-making, even when dealing with major uncertainties and limited foreseeability. It clearly indicates that AM is the best option that we have now. Also, we seem to have good reasons to prefer MM over S-SRM deployment in the here and now. The circumstances for this assessment, however, might change in the future.

Either way, no matter whether one agrees with our analysis of S-SRM in Section III, we hope to have at least shown that employing a satisficing multi-dimensional consequentialist framework in the context of climate change policy options does manage to deal with uncertainties and risk, and provides plausible answers. It allows us to carefully assess people's wellbeing, and an option's probability and fairness, even in complex situations; because of that, we take our satisficing multi-dimensional consequentialist framework to be an attractive assessment tool for ethicists and policymakers alike, helping to reach morally permissible decisions in complex choice situations.

ACKNOWLEDGEMENTS

Research for this paper has benefited from participation in the Research Networking Programme 'Rights to a Green Future', which is financed by the European Science Foundation, as well as from financial support by the UFSP Ethik of the University of Zurich. Work was also supported by the CEMICS (Contextualizing Climate Engineering and Mitigation: Illusion, Complement or Substitute) project of the Priority Programme (SPP 1689) of the German Research Foundation. Earlier versions of the paper were presented at the ECPR Joint Sessions in Mainz, as well as the Institute for Advanced Sustainability Studies (IASS) in Potsdam. We would like to thank the audiences at both occasions for helpful comments and criticisms. In particular we would like to thank Gregor Betz, Tobi Svoboda, Martin Peterson and Simon Hailwood, as well as two anonymous referees for *Environmental Values*, for their insightful suggestions on how to improve the text.

REFERENCES

Bodansky, D. 2011. 'Governing climate engineering: scenarios for analysis discussion'. Paper 2011–47, Harvard Project on Climate Agreements.

Bradley, B. 2006. 'Against satisficing consequentialism'. *Utilitas* **18**: 97–108. http://dx.doi.org/10.1017/S0953820806001877

Burns, W. 2011. 'Climate geoengineering: solar radiation management and its implications for intergenerational equity'. *Stanford Journal of Law, Science & Policy* **4**: 39–55.

Darwall, S. (ed.) 2003. *Consequentialism*. Oxford: Blackwell.

Davies, G. 2010. 'Geoengineering: a critique'. *Climate Law* **1**: 429–441.

Goes, M., N. Tuana and K. Keller. 2011. 'The economics (or lack thereof) of aerosol geoengineering'. *Climatic Change* **109**: 719–744. http://dx.doi.org/10.1007/s10584-010-9961-z

Jackson, F. 1991. 'Decision-theoretic consequentialism and the nearest and dearest objection'. *Ethics* **101**: 461–482. http://dx.doi.org/10.1086/293312

Lenman, J. 2000. 'Consequentialism and cluelessness'. *Philosophy & Public Affairs* **29**: 342–70. http://dx.doi.org/10.1111/j.1088-4963.2000.00342.x

Louis, M.E.S. and J.J. Hess. 2008. 'Climate change impacts on and implications for global health'. *American Journal of Preventive Medicine* **35**: 527–538.

Meyer, L.H. and D. Roser. 2009. 'Enough for the future'. In A. Gosseries and L.H. Meyer (eds), *Intergenerational Justice*, pp. 219–248. Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199282951.003.0009

Meyer, L.H. and D. Roser. 2010. 'Climate justice and historical emissions'. In L.H. Meyer and M. Matravers (eds), *'Justice, Equality and Democracy'*, special issue of *Critical Review of International Social and Political Philosophy* **13**: 229–253. http://dx.doi.org/10.1080/13698230903326349

Mulgan, T. 1993. 'Slote's satisfying consequentialism'. *Ratio* **6**: 121–134. http://dx.doi.org/10.1111/j.1467-9329.1993.tb00142.x

Mulgan, T. 2001. 'How satisficers get away with murder'. *International Journal of Philosophical Studies* **9**: 41–46. http://dx.doi.org/10.1080/09672550010012129

Norcross, A. 1990. 'Consequentialism and the unforeseeable future'. *Analysis* **50**: 253–256. http://dx.doi.org/10.1093/analys/50.4.253

Oberdiek, J. 2012. 'The moral significance of risking'. *Legal Theory* **18**: 339–356. http://dx.doi.org/10.1017/S1352325212000018

Peterson, M. 2012. 'Multi-dimensional consequentialism'. *Ratio* **25**: 177–194. http://dx.doi.org/10.1111/j.1467-9329.2012.00530.x

Peterson, M. 2013. *The Dimensions of Consequentialism: Ethics, Equality and Risk*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139519243

Pettit, P. 1984. 'Satisficing consequentialism'. *Proceedings of the Aristotelian Society*, Supplementary Volumes **58**: 165–176.

Preston, C.J. 2012. 'Solar radiation management and vulnerable populations: The moral deficit and its prospect'. In C.J. Preston (ed.), *Engineering the Climate: The Ethics of Solar Radiation Management*, pp. 77–93. Plymouth: Lexington Books.

Ricke, K.L., M.G. Morgan and M.R. Allen. 2010. 'Regional climate response to solar-radiation management'. *Nature Geoscience* **3**: 537–41. http://dx.doi.org/10.1038/ngeo915

Robock, A., L. Oman and G.L. Stenchikov. 2008. 'Regional climate responses to geoengineering with tropical SO_2 injections'. *Journal of Geophysical Research-Atmospheres* **113**, D16.

Rogers, J. 2010. 'In defence of a version of satisficing consequentialism'. *Utilitas* **22**: 198–221. http://dx.doi.org/10.1017/S0953820810000099

Ross, A. and H. Damon Matthews. 2009. 'Climate engineering and the risk of rapid climate change'. *Environmental Research Letters* **4**: 45–103. http://dx.doi.org/10.1088/1748-9326/4/4/045103

Schneider, S.H., S. Semenov, A. Patwardhan, I. Burton, C.H.D. Magadza, M. Oppenheimer, A.B. Pittock, A. Rahman, J.B. Smith, A. Suarez and F. Yamin. 2007. 'Assessing key vulnerabilities and the risk from climate change'. In M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E.

Hanson (eds), *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 779–810. Cambridge: Cambridge University Press.

Schäfer, S., M. Lawrence, H. Stelzer, W. Born, S. Low, A. Aaheim, P. Adriázola, G. Betz, O. Boucher, A. Carius, P. Devine-Right, A.T. Gullberg, S. Haszeldine, J. Haywood, K. Houghton, R. Ibarrola, P. Irvine, J.-E. Kristjansson, T. Lenton, J.A. Link, A. Maas, L. Meyer, H. Muri, A. Oschlies, A. Proelß, T. Rayner, W. Rickels, L. Ruthner, J. Scheffran, H. Schmidt, M. Schulz, V. Scott, S. Shackley, D. Tänzler, M. Watson, N. Vaughan. 2015. *The European Transdisciplinary Assessment of Climate Engineering (EuTRACE): Removing Greenhouse Gases from the Atmosphere and Reflecting Sunlight away from Earth.* European Union Seventh Framework Programme.

Schuppert, F. 2013. 'Distinguishing basic needs and fundamental interests'. *Critical Review of International Social and Political Philosophy* **16**: 24–44. http://dx.doi.org/10.1080/13698230.2011.583532

Sinnott-Armstrong, W. 2012. 'Consequentialism'. *Stanford Encyclopedia of Philosophy*. Online at: plato.stanford.edu/entries/consequentialism/ (accessed 5 September 2015).

Slote, M. 1984. 'Satisficing consequentialism'. *Proceedings of the Aristotelian Society*, Supplementary Volumes **58**: 139–63.

Smart, J.J.C. and B. Williams. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511840852

Smart, J.J.C. 1973. 'An outline of a system of utilitarian ethics'. In J.J.C. Smart and B. Williams (eds), *Utilitarianism: For and Against*, pp. 3–74. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511840852.001

Svoboda, T., K. Keller, M. Goes, and N. Tuana. 2011. 'Sulfate aerosol geoengineering: The question of justice'. *Public Affairs Quarterly* **25**: 157–179.

Thomson, J.J. 1986. *Rights, Restitution, and Risk*. Cambridge, MA: Harvard University Press.

Tuana, N. R.L. Sriver, T. Svoboda, R. Olson, P.J. Irvine, J. Haqq-Misra and K. Keller. 2012. 'Towards integrated ethical and scientific analysis of geoengineering: A research agenda'. *Ethics, Policy & Environment* **15**: 136–157. http://dx.doi.org/10.1080/21550085.2012.685557

Tuana, N. 2013. 'The ethical dimensions of geoengineering: Solar radiation management through sulphate particle injection'. *Geoengineering Our Climate*, Working Paper and Opinion Article Series. Online at: http://geoengineeringourclimate.com/2013/06/11/the-ethical-dimensions-of-geoengineering-solar-radiation-management-through-sulphate-particle-injection-working-paper/ (accessed 5 September 2015).