

# *ETHICS AND RISKS: APPROACHES AND ISSUES*

## INTRODUCTION

Most of our decisions are made under conditions of risk and/or uncertainty about the possible consequences of our actions or omissions. Examples range from mundane choices (e.g. driving or not driving to work, getting an influenza vaccine or not) to far-reaching political decisions (e.g. about climate change, terrorist attacks or new and emerging technologies). Given the inescapability and significance of risk and uncertainty, it is little wonder that ethicists' concern with these topics has been increasing at a steady pace.

The first challenge when engaging with the philosophy of risk and uncertainty is terminological. Depending on context, the term 'risk' admits of many uses and definitions (e.g. risk may denote an unwanted event, the probability of an unwanted event or the cause of an unwanted event). Nevertheless, commonalities exist. What most usages of the term have in common is that they regard risks as something undesirable that involves a lack of knowledge about what will happen.

A widespread, albeit contentious, starting point in the current philosophical debate is to conceptualize risk as a combination of the severity and the probability of occurrence of an unwanted event. Such cases of action under risk are distinguished from cases of action under uncertainty, in which we lack probability information for the unwanted events, and cases of deep uncertainty, in which even the possible consequences of an action are unknown.

Many of the risks we face today result from technological and societal advances, and are imposed on us by other agents and/or collectives. As such, not only is risk imposition itself an appropriate subject of normative evaluation, but also its management, (re)distribution, and reparation. Furthermore, additional questions loom in the background, such as what a

justifiable overall level of risk consists in and how to handle existential risk.

Classical ethical theories rarely address risk explicitly, and inquiry tends to focus disproportionately on actions with determined outcomes. There may be two reasons for this oversight. First, the need to keep discussion and examples tidy and free from complicating factors. Second, the assumption of a division of labour between ethics and decision theory, in which ethics tells us what to do and decision theory tells us how to implement these answers when faced with risk and uncertainty. Problems are exacerbated by the fact that even if theoretical or ideal-level responses to normative questions raised by risk were to be found, these would have to be translated into action-guiding recommendations in a non-ideal world. Needless to say, such a task would not be an easy one.

However, burying our heads in the sand is not the solution. Risk imposition is ubiquitous and introduces genuine ethical dilemmas. This special issue not only demonstrates this fact, but it goes a considerable distance in tackling problems like those mentioned above, as well as other risk-related difficulties. It brings together authors with divergent theoretical commitments and different perspectives around a multi-faceted problem: squaring ethics with risks. The contributions collected here reflect the complexity of the subject matter and its wide-ranging importance.

Friedemann Bieber argues that philosophers have not devoted enough energy to the development of a conceptual framework for distinguishing different types of risks. He first illustrates the need for a more nuanced conceptual framework. This is based on the thesis that ethicists need to distinguish between different types of risks because of the great variety in risk-related normative considerations. Subsequently, the author develops two ideas. First, he proposes that we understand ‘risk’ as a multi-dimensional concept – a concept that can always be evaluated in different dimensions (e.g. a qualitative, a quantitative and an epistemic dimension). Second, he introduces an original account of ‘systemic risk’. Criticising current accounts, Bieber suggests we conceptualize ‘systemic

risk' as a risk that endangers the functioning of a system and originates in the system or its background noise. Further, the author identifies a number of structural features shared by systemic risks and evaluates the distinctive normative considerations to which these give rise. In a final step, Bieber argues that the notion of 'systemic risk' is particularly apt for characterising some of the most pressing risks we face in our interconnected world.

Sven Ove Hansson takes the fact that risk-related decisions are often hindered by conflicts and uncertainties at the science-policy interface as his starting point. The author shows that there is ample room for improving decision-making processes and their employment of science. Hansson argues that philosophical tools and distinctions can have a vital role in eliminating or mitigating risk-related uncertainties. The author advances the position that part of the reason why philosophy (in particular, philosophy of science, decision theory, and ethics) can have such a role is because many of the so-called scientific controversies are not conflicts within science, but between science and science denial or forms of pseudoscience. In such cases, alleged uncertainties can be overcome by doing away with pseudoscientific claims and fabricated controversies. The author argues that a significant part of any remaining uncertainty will not be relevant to decisions, which he demonstrates *via* two simple symmetry tests: the test of alternative causes and the test of opposite effects. In a further step, Hansson argues that decision theory and ethics offer ways of handling remaining decision-relevant scientific uncertainties, such as hypothetical retrospection. The latter consists in a systematic deliberation on how we will evaluate present decisions in different potential future situations. The purpose of this process is to approach the ideal of making ethically defensible decisions regardless of what happens.

Madeleine Hayenhjelm starts by presenting a host of puzzling questions. What kind of moral principle could strike the right balance between being sufficiently restrictive to avoid the kind of large-scale risks that have resulted in catastrophe in the past, but not stifle desirable progress?

Is there such a principle that is not merely a precautionary principle, but one that could be based on firm moral grounds? Hayenhjelm sets out to explore a deceptively simple thought, namely that *reparability* serves as a moral constraint against risky policy decisions. Her idea comes in two forms. First, it is morally wrong to impose a risk of harm that is, in principle, irreparable, such that it would bring about a permanent loss of a kind *qua* kind. Second, it is morally wrong to impose a risk of reparable (but not compensable) harm that exceeds what could be repaired in actuality. She aims to achieve two things. First, she sets out to describe the moral problem to be addressed for any principle-guiding decisions about risk. This involves acknowledging that risk decisions are epistemically impaired decisions, and that both epistemic and moral uncertainty must be taken into account alongside outcome uncertainty. Second, Hayenhjelm introduces the idea of reparability as a moral constraint in the form of two versions of the Reparability Principle. In conclusion, the author argues that the Reparability Principle is both a morally intuitive and valuable guide to decision-making.

Lukas H. Meyer and Harald Stelzer argue that climate change-induced uncertainties raise significant challenges to normative theory. One such challenge arises because currently living people's actions do not necessarily harm future people, but they do impose risks of rights-violations onto them. This gives rise to important questions. How should we weigh the climate change induced impositions of risks on future people? How can we comparatively assess the imposition of these risks to the interests of people currently living and the costs, harms and forgone benefits by climate change policies? The authors advance a framework for answering some of these questions by outlining a risk-averse interpretation of intergenerational sufficientarianism. They argue in favour of this position and reveal its implications for climate change responses.

Klaus Steigleder addresses the oft-made complaint that rights-based moral theories cannot adequately or convincingly deal with risks. Broadly speaking, rights-based moral theories prohibit the imposition of risks that

cannot be adequately controlled and that involve the possibility of basic harms (e.g. death). Nonetheless, circumstances exist in which the imposition of such risks is permitted or even required. Steigleder attempts to show when this is the case by advancing a non-exhaustive list of five criteria. First, the imposition of a mere risk of a basic harm may be preferable to its secure infliction when certain conditions obtain. Second, the imposition of the risk of a basic harm on one person may be justified when this is the only way in which a risk of an even greater harm to another person can be prevented. Third, when the imposition of a small-probability risk of grave harm on one person is the only way in which a high-probability risk of grave harm to another person can be prevented, then this may be situationally justified. Fourth, if the normative inescapability of basic harm risks pertains to risks that are typically connected to certain technologies, technological systems and forms of social organization, then such risks can be justified if a technology or form of organization helps to avert even greater risks to all affected. Fifth and finally, the imposition of risks of basic harm can be justified through the consent of those affected.

Jeffrey Haines advances three arguments. First, political libertarians should hold that it is wrong to create a risk of non-negligible harm for non-consenting others. Second, compensatory payments have to be made to those put at risk in order to secure their consent to such risk imposition. Third, given that such a system of payments would be highly impractical and would require a massively invasive bureaucracy to enforce, libertarians should prefer a generalized system of redistribution as an alternative. Such a system would base receipt of payments not on need but on exposure to risk. The system would require larger payments from those who create more risk for other people and not, necessarily, from those who are wealthier. Nevertheless, to the extent that wealthier people create more risk for the poor than vice versa (e.g. the driver of the car creates a bigger risk for the pedestrian than the pedestrian does for the driver of the car), the pattern of redistributive payments will resemble that

of conventional theories of distributive justice. Provided the correctness of his account, Haines concludes that considerations of risk imply that libertarians should support a fairly robust redistributive state.

Nimrod Z. Kovner argues that putting others at risk can establish morally corrupt relations between individuals. Kovner's argument tracks and advances the legal and moral philosophy literature on the moral significance of risk in normative thinking. The author claims that understanding why imposing risk bears moral significance is vital for answering questions regarding the meaning of liability or redress in the context of risk-imposition. Scholars have tended to focus on the impact of risk impositions on individuals' mental state, resources, interests or autonomy. Kovner's primary contribution lies in his emphasis on the relational aspects of risk-imposition. The author claims that putting others at higher risk may form a morally defective relation. He advances this view by way of an analysis of two potentially defective relations: one of domination, which is based on Pettit's work on Republican freedom and one of exploitation, which occurs when an agent takes unfair advantage of another agent's vulnerability.

Aviezer Tucker argues that justice in the distribution of risks, risk assessments, and trust in expertise, technocratic and political elites are closely linked. Distributive justice is not solely a normative theory that should regulate the distributions of risks and rewards, but also a political force that affects policies. Tucker begins by outlining David Schmidtz' theory of justice, as well as the social theories of risk of Beck, Luhmann and Giddens. In a subsequent step, he describes the risks that accompany new technologies of extracting unconventional energy resources. Tucker also outlines policy responses to the implementation of these technologies in Europe and applies the theories expounded at the outset to explain the political reactions and policy results in Europe. The author highlights a feedback causal loop between unaccountable technocratic and political elites and distrust. He asserts that geopolitical risk may balance technological risk assessment when assuming technological risks reduce geopolitical

## INTRODUCTION

risks. In a final step, Tucker's account is buttressed by a comparison of regulatory policies for the extraction of unconventional energy resources in the US, as well as an analysis of a case in England where the government attempted to decrease injustice in the distribution of risks.

Adriana Placani, Eike Düvel, and Thomas Pözlner

# *Addressing the Diversity of Risks and Accounting for Systemic Risks: Two Proposals for Improving Clarity in Philosophical Discussions of Risk*

Friedemann Bieber  
University of Zurich, Switzerland

**ABSTRACT.** The lack of engagement of philosophy with decisions made under conditions of risk and uncertainty has lately received increasing attention. But philosophers have devoted little thought to the development of a conceptual framework for distinguishing different types of risks. This article begins by illustrating the need for a more nuanced conceptual framework. As the normative considerations risks give rise to are highly varied, ethicists need to distinguish between different types of risks. It then offers two ideas. First, it proposes that we understand ‘risk’ as a multi-dimensional concept, i.e. a concept that can always be evaluated in different dimensions, e.g. a qualitative, a quantitative and an epistemic dimension. While this understanding of risk underlies the current practice of risk analysis, making it explicit is of value, in particular for the philosophical debate. It draws attention to the diversity of risks, helps to group them sensibly and sharpens existing arguments. Second, the article introduces a novel account of the notion ‘systemic risk’. Criticising existing accounts, it proposes we understand ‘systemic risk’ as referring to a risk that endangers the functioning of a system and originates in the system or its background noise, where a system is conceived of as a set of interdependent, interacting components that form a complex, functionally defined whole. The article identifies a number of structural features shared by systemic risks and evaluates the distinctive normative considerations they give rise to. It then argues that the notion ‘systemic risk’ is particularly suited to characterise some of the most pressing risks faced in an increasingly interconnected world. While the notion ‘systemic risk’ rests on the understanding of ‘risk’ as a multi-dimensional concept, it also complements it. Its necessity shows that in addressing risks, it is sometimes helpful to move beyond a characterisation based solely on their dimensions.

**KEYWORDS.** Risk, systemic risk, uncertainty, conceptual framework

## I. THE DIVERSITY OF RISKS AND WHY IT MATTERS

In normative ethics, philosophers often entertain hypothetical decision scenarios which assume that there is a binary choice and that for each of the two options the resulting outcome is known with certainty. This is illustrated, for instance, by renowned thought experiments like the Trolley Problem (Foot 1967), Jim and the Indians (Williams 1985) and Ticking Bomb (Shue 1978). A criticism that has lately been raised is that the assumption of certainty of outcomes amounts to a problematic simplification (see, for example, Fried 2012; Hansson 2013).<sup>1</sup> It is a simplification, it is said, because when actually making decisions, one ordinarily lacks knowledge of the outcome that would result from each possible choice. This simplification is problematic, it is further said, because it ignores the complexities arising from making and justifying decisions under conditions of risk. While this diagnosis has been followed by a number of publications addressing the topic of risk, either by invoking risk as a resource to challenge specific theories in normative ethics or by directly addressing the question of how one should act in a certain situation involving a risk, philosophers have – with some exceptions, notably Hansson (e.g. 2013) – given comparatively little thought to the development of a comprehensive conceptual framework for grouping and classifying different types of risk.

But risks are highly diverse. There are risks that affect individuals (the risk of being hit by lightning) and risks that affect collectives (the risk of flooding faced by an island population); there are risks that mainly affect currently living individuals (the risk of an earthquake) and risks that primarily affect future individuals (the risks of climate change); there are risks where harm is reversible (the risk of losing money in a bet) and risks where harm is irreversible (the risk of being killed). Risks of each such type give rise to different normative considerations. For example, whether or not a potential harm is reversible will affect what will be considered an adequate justification for imposing the risk. It seems *prima facie* plausible

that you can justify the imposition of a risk of inadvertently damaging my bicycle while driving on the grounds that driving is of great benefit to you and that you will compensate me if the risk materialises. But you need a more sophisticated justification for the imposition of a risk of killing me while driving, because the harm threatened is irreversible and you could therefore not provide adequate compensation if the risk were to materialise. One strategy for justifying the imposition of this risk consists in pointing out that the practice of driving is beneficial to society at large and that the risks resulting from this practice are imposed in a broadly symmetrical pattern. Whether or not this strategy is successful, the need for a different justification in the case of irreversible harm illustrates the more general point. The risks we face vary and the normative considerations we need to take into account diverge accordingly. Without an adequate conceptual framework, we are unable to identify and address them effectively. So, developing such a framework needs to be a fundamental concern to the ethicist who wishes to enquire how we should act under conditions of risk.

This article aims to contribute to the development of such a conceptual framework in two distinct ways. First, by drawing on the practice of risk studies and ordinary discourse, it develops a proposal for understanding ‘risk’ as a multi-dimensional concept, taking a multi-dimensional concept to be a concept that can always be evaluated in different dimensions, where the relevance of specific dimensions is context-sensitive. Understanding ‘risk’ in this way draws attention to the diversity of risks, can help avoid misunderstandings and will permit improvements in the classification of risks, thereby contributing to the development of more nuanced arguments on how to handle them. Second, and building on this proposal, the article introduces a novel account of the notion ‘systemic risk’. It argues that an important and distinctive type of risk exists that arises from a specific form of systemic interaction, but has so far been largely overlooked by philosophers. While a few accounts of the notion ‘systemic risk’ have been introduced, none of them, it is argued, captures

the characteristics of this type of risk adequately. Based on the novel account proposed, the article explores the structural features of systemic risks and the distinctive normative considerations to which they give rise. In so doing, it provides an argument for why philosophers should pay more attention to such risks.

The two parts of the article are connected. On the one hand, the proposal of a particular understanding of the notion ‘systemic risk’ builds on the more general proposal of understanding ‘risk’ as a multi-dimensional concept. On the other hand, it complements it, showing that for the purpose of a clear conceptual framework it is sometimes necessary to classify risks not merely according to their dimensions, but according to their underlying structural features.

## II. UNDERSTANDING ‘RISK’ AS A MULTI-DIMENSIONAL CONCEPT

The term ‘risk’ is ambiguous. The sentences “Crossing a red light involves the risk of being hit by a car” and “There is a five per-cent annual risk of flooding” both sound natural, but invoke distinct notions of risk. In the first sentence, ‘risk’ refers to an undesirable outcome that may materialise; in the second sentence, ‘risk’ refers to the likelihood of such an undesirable outcome. Hansson (2014, 2-3) distinguishes five fundamental senses in which the word ‘risk’ is used: as referring to (i) an unwanted event that may occur, (ii) the cause of an unwanted event that may occur, (iii) the probability of an unwanted event that may occur, (iv) the statistical expectation value of an unwanted event that may occur, (v) the fact that a decision is made under the condition that the probabilities are known.<sup>2</sup>

In philosophical debate, this ambiguity is frequently disregarded. For instance, many recent articles (see, for example, Frick 2015) that address risk in normative ethics effectively employ the term ‘risk’ in its decision-theoretic sense (v), discussing only cases in which all probabilities are known. They thereby rely on a distinction first made by Knight (1935 [1921]), according to which a decision is made under conditions of risk

rather than uncertainty when all probabilities are known. Drawing on a classification schema for types of uncertainty proposed by Bradley and Drechsler (2014), a situation of risk could be characterised as one involving only mild uncertainty: we are not certain what the true state of the actual world is, but we know all feasible actions, all possible consequences, and all possible states of the world (no state-space uncertainty) as well as the consequences of each action in each state of the world (no option uncertainty). But it is clear that the decisions we confront in real life vary in the extent to which we have knowledge of available options, possible outcomes and conditional probabilities.<sup>3</sup> In ordinary discourse, however, decisions made in the absence of knowledge of the exact probabilities are still said to be made under conditions of risk, or to involve risk. So, limiting consideration to a subset of cases of risk while not making this explicit may be misleading – talking about risks quite generally can suggest a wider scope for the theses stated than is in fact justified by the arguments.

### *The Concept of Risk as a Multi-Dimensional Concept*

I here propose to understand the concept of risk as a multi-dimensional concept. I take a multi-dimensional concept to be a concept that always can and often needs to be evaluated in different dimensions (or according to different measures), where the relevance of a particular dimension (or measurement) varies across contexts; as a result, in employing a multi-dimensional concept, one has to reflect on and make explicit the dimension(s) to which one relates. Of course, the idea that risks can be evaluated in different dimensions is not new; a distinction of dimensions is standard practice in the field of risk analysis. I argue, nevertheless, that it is valuable to make this idea explicit, in particular for debates in philosophy. In the following, I proceed by first sketching my proposal, then relating it to the existing literature and practice and finally examining what renders it fruitful.

As was illustrated above, the term ‘risk’ is ambiguous. In some technical contexts, for instance in decision theory, it is commonly used only

in one specific sense. Yet, ‘risk’ is not, generally, a technical term, but a term from ordinary language. So, at least outside the realm of some very restricted discourse among experts, we cannot simply redefine the term in a way that suits our needs – doing so would give rise to misunderstandings; moreover, it might result in an evasion of the questions that are actually of interest, which are tied to the ordinary understanding of risk. So, the moral philosopher will have to rely on a notion of risk that is compatible with ordinary language. As empirical studies show, the term ‘risk’ is ordinarily used in a great variety of ways, though rarely in its more technical senses (Boholm, Möller and Hansson 2016). What these uses have in common is that they in some way refer, at the most general level, to the possibility that an undesirable event will occur. Moreover, as Hansson (2013, 11) illustrates, the concept of risk is both value-laden and fact-laden. It is value-laden because ‘risk’ always refers to the possibility of the occurrence of an event that is undesirable, so in saying that something is a risk one expresses the value-judgement that the occurrence of some event would be undesirable. It is fact-laden, because in saying that a risk exists, one claims that it is (for all one knows) factually possible that an unwanted event will occur.

When confronting any risk, understood most broadly as the possibility of an undesirable event, there are three fundamental questions one needs to consider. How undesirable is the possible event? How likely is it to occur? How strong are the epistemic grounds on which these answers rest? These three questions refer to three dimensions, in which any risk can be evaluated: a qualitative dimension (the extent of the potential harm), a quantitative dimension (the likelihood of the potential harm) and an epistemic dimension (the confidence one has in the answers to the first two questions). In a way, these dimensions are alluded to by the five senses of the word ‘risk’ identified above, which can be divided into three groups. (i) and (ii) refer to the possible unwanted event or its source; they allude to the qualitative dimension. (iii) and (iv) refer to some measure of the likelihood that an undesirable event occurs; they allude to the

quantitative dimension. (v) refers to a specific form of knowledge; it alludes to the epistemic dimension. So, while the word ‘risk’ is used in different, specific ways, these can all be accounted for by the more general understanding of risk as the possibility of an undesirable event, if one allows for a risk to be evaluated in different dimensions.

On the basis of this examination, I propose that we understand ‘risk’ as a multi-dimensional concept. By this, I mean a concept that can and frequently must be evaluated in different dimensions. In practice, this means that in characterising a risk, one will need to describe it with regard to different dimensions; and in comparing risks, one will need to be explicit about the dimensions one is concerned with. If I ask my doctor to characterise the risk involved in undergoing a certain operation, she will need to describe it with regard to different dimensions, stating, for example, what she considers to be the set of undesirable outcomes that might occur, their respective probabilities and her evidence for these estimates. And if I ask her to compare it to the risk involved in an alternative procedure, she has to evaluate them separately in each dimension or, if she provides a single verdict, be explicit about how she weighed the dimensions.

In proposing that we understand ‘risk’ as a multi-dimensional concept, I consider the quantitative, qualitative and epistemic dimensions as fundamental, but not exhaustive. At least in some contexts, additional dimensions will be relevant. For instance, whether the effects of climate change are reversible appears to be critical for the appropriate level of precaution, so the dimension of reversibility will be important in describing the risk of climate change. Similarly, a dimension of time-lag (capturing the time-horizon in which a risk may materialise), a dimension of distribution (capturing how equally people would be affected) or a dimension of persistence (capturing the durability of possible effects) – to name just a few – will become relevant in some contexts. However, these further dimensions can be ignored in the context of simpler risks. For example, in evaluating the risk involved in crossing a street, reversibility, time

lag, distribution and persistence of the possible harm typically do not warrant special consideration. What we are concerned with can be summarised by three measures. How undesirable the outcome would be, how likely it is and how good the epistemic position is from which one makes these judgements.<sup>4</sup>

### *Locating the Proposal in Literature and Practice*

Of course, the idea that there are different dimensions in which a risk can be evaluated is not new. In a sense, this is how we always evaluate risks, although we might not typically make this explicit. Moreover, Hansson (1989), looking at risks regulated at the political level, already distinguishes eight dimensions in which risks can be assessed. And in the field of risk studies, an explicit distinction of different dimensions in the description of risks has long been standard practice, not least because it allows one to group risks in useful ways. Perhaps the most impressive illustration of this is a report by the German Advisory Council on Global Change (WBGU) on future risks faced by societies (Schellnhuber *et al.* 1998). The report initially distinguishes nine criteria according to which risks may be compared and evaluated. Borrowing labels from Greek mythology, it then identifies a number of relevant classes of risks by reference to these criteria. For example, risks of the *Sword of Damocles* type are characterised by highly undesirable outcomes, a low probability of materialisation and a high confidence in these estimates (Schellnhuber *et al.* 1998, 11) – the risk of an impact of a large meteorite falls into this class. So, *Sword of Damocles* risks are characterised by a combination of particular features in each of the three dimensions I identified above as fundamental. But the classificatory scheme of the WBGU moves beyond the qualitative, quantitative and epistemic dimension. For example, risks of the *Pandora* type are characterised not only by great uncertainty about probability and outcomes, but also by the high persistence and limited reversibility of their potential effects (Schellnhuber *et al.* 1998, 11-12). The approach taken by the

WBGU can therefore be said to be an example of the understanding of risk here advocated – to take certain dimensions of risk as fundamental and to add additional dimensions as needed.

But while evaluating risks in different dimensions has long been common practice, most definitions of the term ‘risk’ do not refer to dimensions. This is evident, for instance, from the overview of common definitions found in Aven and Renn (2009, 1-2). Still, some such definitions are in principle compatible with the understanding of the concept of risk here advanced. Rosa, for instance, proposes the following definition: “Risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain” (1998, 28). Setting aside all other concerns with this definition, it may be said to be compatible with the idea that there are different dimensions to risk. We could assess a situation or an event according to what is at stake, how likely it is that what is at stake will be lost and how confident we are about these estimates. But we could also move beyond the qualitative, quantitative and epistemic dimension, assessing, for instance, the extent to which what is at stake might later be recovered (dimension of reversibility). So, the general understanding of risk as a multi-dimensional concept is compatible with specific definitions of risk.

Some accounts of risk do explicitly introduce the idea of dimensions, but not in the way understood here. Aven, for instance, states that “[...] risk has two main dimensions – consequences and uncertainties” (2015, 15). While Aven does not provide a clear definition of risk, he apparently takes a risk to consist in an ordered pair, a set of consequences and a set of uncertainties.<sup>5</sup> In the examples he provides, the statement of a risk consists of two components, the statement of a set of consequences, such as the occurrence (or non-occurrence) of specific events, and the statements of a set of things currently unknown, such as whether or when a specific event will occur. He contrasts risks with risk descriptions. The risk descriptions differ, roughly speaking, by including specific measures of the likelihood of the relevant outcomes, which are given conditional

on the models, data and further assumptions (labelled ‘background knowledge’) in the light of which the risk is described. Such an approach might be suited to risk analysis because it allows one to evaluate how variations in the models affect the conditional predictions of outcomes. But this comes at a cost. Most importantly, Aven’s notion of risk is technical and conflicts with the ordinary understanding. We simply do not think of risks as ordered pairs. Given that we always rely on some background knowledge, it seems that we would, on Aven’s account, only be having discussions about risk descriptions, not the risks themselves, which would exist in some ontological realm we could not reach. This also explains why a risk, on Aven’s account, has only two dimensions, whereas a risk description could be said to have three. Introducing some probability estimates adds a quantitative dimension. But in ordinary language, we do not use the word ‘risk’ in the way Aven does; we evaluate and compare risks rather than merely risk descriptions. So, I take it, my proposal of distinguishing different dimensions of risk is more plausible.

To summarise, the proposal of understanding ‘risk’ as a multi-dimensional concept cannot lay claim to great originality. Such an understanding of risk can be said to underpin our ordinary thinking about risk. But since I do not intend to propose a technical definition of risk, but to provide an account of risk that is in line with common understanding, this is only welcomed. Nonetheless, proposing that we understand the concept of risk as a multi-dimensional concept is not trivial. It makes explicit what usually remains implicit and is often overlooked in theoretical discussions. As the next section shows, the proposal can be conducive to improving clarity, in particular in philosophical debate.

### *Advantages of this Understanding of Risk for Philosophical Debate*

Understanding ‘risk’ as a multi-dimensional concept, and being explicit about this, can be fruitful to philosophical debate in several ways. First, it can help avoid misunderstandings and render arguments more precise.

As noted at the beginning of this section, philosophers do not always reflect on the ambiguity of the term ‘risk’, at times using the word in a narrow or technical sense without making this explicit. This, I argue, is problematic because it obscures the scope of the argument. As Hansson notes, the technical use of the term ‘risk’ has led to “significant communicative failures” (2013, 10). In Frick (2015), which was introduced as an example, all cases of risk discussed are cases of known probabilities. But it cannot be assumed, without argument, that the same analysis applies to cases involving considerable uncertainty – the extent of one’s knowledge might make a difference. Now, if one were to understand ‘risk’ as a multi-dimensional concept, one would need to specify the risks under consideration further, characterising them in each relevant dimension. But doing so would make explicit that the risks actually considered constitute, at least in the epistemic dimension, a limiting case. In doing so, one would also be forced to address the question whether or not this restriction is problematic. So, understanding ‘risk’ as a multi-dimensional concept would ideally not only help avoid misunderstandings, but also render arguments more precise.

Second, this understanding of risk could help make disagreements in value judgments apparent. In comparing risks, we typically reach one overall verdict. For instance, we might say the risk of a car accident when driving to place A is greater than the risk of a plane crash when flying to place A. It seems, then, that we rely on a single measure for the severity of the risk. But there might be reasonable disagreement about how this single measure should be composed. Because risk is a value-laden notion, a bigger risk is *ceteris paribus* worse than a smaller risk. But it is possible that two people disagree about whether the risk of car accident is greater than the risk of plane crash while agreeing in the estimates of the undesirableness of each event, its likelihood and the reliability of these estimates. In this case, their disagreement would be based in a divergent weighing of the dimensions of risk. For instance, one person might attach greater weight to the qualitative dimension and, reasoning that the

outcome of a plane crash would be more undesirable, judge flying to involve the greater risk. The other person might attach greater weight to the quantitative dimension and, reasoning that the likelihood of a car accident is higher, judge driving to involve the greater risk. So, taking risk to have different dimensions might not only render transparent the origins of disagreement about the ranking of risks, but might also prompt arguments about how one should weigh different dimensions in a given context.

Third, understanding ‘risk’ as a multi-dimensional concept can be of help in grouping and classifying risks in a sensible manner. As was argued above, different risks give rise to distinct normative concerns. The imposition of a risk of irreversible harm will in general require a different justification than the imposition of a risk of reversible harm. But risks that are similar in all relevant dimensions can be expected to give rise to similar normative concerns, so classifying them by reference to different dimensions can be helpful. The WBGU’s proposal could serve as a role-model (and illustrates how philosophers can benefit from taking into account literature on the topic of risk published in other subjects). It provides a framework for grouping risks in a nuanced way, which is essential for developing strategies aimed at effectively responding to them. Of course, the WBGU has a different primary objective than the moral philosopher. It is primarily interested in better understanding the risks societies face (e.g. emphasising certain areas for scientific research), in rising awareness for these risks in the public sphere (e.g. through risk communication) and in developing new strategies for risk-management (e.g. alternatives to trial-and-error methods for risks of irreversible harm). But just as the more nuanced conceptual framework helped the WBGU in developing tailored approaches, so it can help the moral philosopher, providing the basis for a more targeted normative enquiry. In some areas of ethics, risks are already classified in a similar way. The debate about existential risks for instance focuses on risks that are distinguished by a particular undesirability and irreversibility of the event that may occur

(cf., for example, Bostrom and Cirkovic 2011). The idea would be to introduce this approach more generally and explicitly.

### III. THE NOTION ‘SYSTEMIC RISK’

In the first part of the present contribution, I proposed that we understand ‘risk’ as a multidimensional concept. I will now develop a novel account of the notion ‘systemic risk’. I begin by arguing that the existing accounts of systemic risk are not adequate. I then develop my own account of systemic risk. I show that systemic risks share certain structural features and give rise to particular normative considerations. Finally, I argue that having a notion of systemic risk is important for the philosophical debate because it sets apart a class of cases of risk that are inadequately captured by the dimensional approach to risk developed above, but of great practical importance.

In the wake of the 2007-2008 financial crisis, ‘systemic risk’ has become a buzzword frequently used in public debate. Although the term can be traced back at least to the 1980s (Cline 1984), its use was initially confined to financial markets. It was only in 2003 that the influential OECD report “Emerging Systemic Risks in the 21<sup>st</sup> Century” (2003) introduced ‘systemic risk’ as a more general term into the wider academic discourse (cf. Renn 2014, 157). It has since been applied in the social sciences (Renn and Klinke 2004; Renn and Keil 2008), in particular in economics (cf. Silva, Kimura and Sobreiro 2017) and has gained a small foothold in philosophy (Heinemann 2014; James 2017; Linarelli 2017).

Nevertheless, as Fouque and Langsam point out, there “[...] is no agreed upon definition of systemic risk” (2013, xxi). Even texts that explicitly introduce the term do not always offer a concise definition. The OECD study exemplifies this. While it contains a section entitled “What are emerging systemic risks?”, this section merely observes that many conventional risks “[...] look set to take on new forms” in the future in a way that will increase the “[...] probability of major vital systems (tech-

nological, infrastructural, ecological, etc.) being severely damaged by a single catastrophic event (natural or man-made), or a complex chain of events” (OECD 2003, 32), thereby suggesting that a systemic risk is to be understood as the risk that a major vital system is severely damaged.

Even where explicit definitions are provided, they tend to be mutually inconsistent. I will illustrate this by reference to Linarelli (2017) and James (2017), who introduce the term ‘systemic risk’ to the philosophical debate, as well as by reference to Renn *et al.* (2017), who develop the most sophisticated account of systemic risk presently found in the field of risk studies. Linarelli, who discusses the demands of justice in the context of systemic financial risk, invokes a definition originally provided by Kaufman, who defines systemic risk as the “[...] the risk or probability of breakdowns (losses) in an entire system as opposed to breakdowns in individual parts or components” (Kaufman 2000, 92; cf. Linarelli 2017, 331). James, who provides an argument for the distinctive significance of systemic risks, states three criteria a risk has to meet in order to qualify as systemic, namely that it is created by “[...] a group of agents [that] act in a coordinated way”, that “[...] the agents’ actions, taken together, suffice to significantly raise the chance that someone or other will suffer serious material injury” and that “[...] no single act or single agent’s actions, taken separately, significantly changes [sic] the probability that harm will occur” (James 2017, 240).<sup>6</sup> Renn *et al.* develop their account by analysing the paradigmatic case of systemic financial risk; they do not offer a definition, but claim that systemic risks have particular properties. They name four descriptive properties – the risks are global in nature, rely on complex causal structures, have non-linear cause-effect relationships and are stochastic in their effect structure – as well as one explanatory property – “[...] systemic risks are due to the interplay of individual micro- and global macro-processes within the system under consideration, combined with exogenous processes that modify the internal dynamics of the system” (Renn *et al.* 2017, 3). These accounts clearly differ: Linarelli does not place restrictions on the origin of the risk; James

does not place restrictions on the extent to which the materialisation of a risk affects the system; Renn *et al.* require risks to be global. As a result, there are at least some risks that only some authors classify as systemic, so their accounts are mutually inconsistent.

### *The Inadequacy of Existing Accounts of Systemic Risk*

In arguing that the existing accounts of systemic risk are inadequate, I limit consideration to the three accounts introduced above. While there are some other proposals (e.g. Schwarcz 2008), they are typically more specific, being concerned exclusively with systemic financial risk. Moreover, for the purpose of this paper, the accounts by Linarelli/Kaufman, James and Renn *et al.* are most significant because they are directly aimed at the philosophical debate and/or most sophisticated in general.

However, before doing so, a word about methodology is in order. In contrast to ‘risk’, ‘systemic risk’ is a technical term, and one that has only recently been introduced. So, the fact that a particular account of systemic risk conflicts with ordinary use does not amount to a decisive objection here. Because ‘systemic risk’ is a technical term, it is in principle permissible to redefine it. But what are the criteria for a good definition of a technical term? While I cannot discuss this question here in detail,<sup>7</sup> I think there are primarily three criteria. First, that the definition is useful with regard to the purpose for which it is introduced (internal adequacy); second, that it can be consistently applied across all fields that have a justified interest in using the term (external adequacy); and third, that the terms constituting the technical term are used in a way that makes sense to the ordinary speaker (semantic adequacy).

The definition invoked by Linarelli is inadequate because it does not place any limits on the origins of the risks that are to be counted as systemic. On his account, the possibility of any event that would result in the breakdown of an entire system qualifies as a systemic risk. But this is implausible. Consider, for instance, the impact of a large-scale meteorite

that devastates an entire region. The impact of such a meteorite might result in the breakdown of all kinds of systems, for instance an electricity supply system or the (regional) financial system. But it would be unhelpful to consider the risk of a meteorite impact a systemic risk (at least with regard to these systems). The reason is that the meteorite's impact has an effect that is completely unrelated to the functioning of these systems. But because it is entirely external to them, the risk of a meteorite impact lacks the characteristic features of the cases typically discussed as systemic risks. This is evident from Linarelli's own paper, which focuses on features such as a complex causal structure that renders it difficult to allocate (causal) responsibility. So, the definition can be said to violate the internal adequacy condition.

The definition introduced by James is inadequate for a similar reason. In not placing restrictions on the forms of material injury, it qualifies risks as systemic that are highly local and confined (relative to the system from which they arise) and therefore lack the features taken as characteristic of systemic risks. For instance, the risk of being harmed in a traffic accident would, on James' account, qualify as a systemic risk. It arises from a coordinated interaction among human agents, this interaction creates the possibility of serious material injury and no single agent's action will significantly affect the level of risk. Yet, it is unhelpful to classify the risk of being harmed in a traffic accident as systemic, because it does not involve the feedback-loops and causal complexity that James himself later focuses on in discussing systemic risks. So, James' definition also violates the internal adequacy condition.

The account of systemic risk advocated by Renn *et al.* is more sophisticated. It limits the scope of the origins and the scope of the potential effects of systemic risks. Still, the account has two deficiencies. The first deficiency is that, echoing the OECD, Renn *et al.* require systemic risks to be "global in nature" (2017, 3). This requirement may be well-motivated by the needs of risk analysts, as risks that endanger systems that are vital at the global level warrant specific attention. But we

ordinarily accept that there are non-global systems and that such systems can give rise to risks. Moreover, these risks may have all the other properties identified by Renn *et al.*; so, requiring systemic risks to be global conflicts with the ordinary use of the term ‘systemic’ and violates the criterion of semantic adequacy.<sup>8</sup> This problem could easily be fixed by dropping the condition that risks are global in nature and introducing the category ‘global systemic risks’ to refer to the relevant subset of systemic risks.<sup>9</sup> A more substantial worry concerns the structure of their account. Renn *et al.* do not present a definition, but propose to understand ‘systemic risk’ as a family resemblance concept,<sup>10</sup> which they characterise by analysing the paradigmatic case of systemic financial risk. One worry is that only considering one paradigmatic example might result in an understanding of systemic risk that is too narrow, because it accepts as definitional certain features of the paradigmatic case; this is illustrated by the above-discussed restriction that systemic risks must be global in nature. More fundamentally, it is questionable whether it is sensible to understand ‘systemic risk’ as a family resemblance concept and to abstain from providing a definition. Because ‘systemic risk’ is a technical term that does not yet have a well-established meaning, it seems plausible that a definition could be given and that it could be useful in further specifying the use of the term.

### *Proposal for a Novel Account of Systemic Risk*

Contrary to the approach taken by Renn *et al.*, who start from a paradigmatic case, I begin at the abstract, conceptual level and attempt to arrive at a proper definition. In so doing, I propose that we understand the term ‘systemic risk’ as referring, most generally, to risks that endanger the functioning of a system and at the same time originate in the system or its background noise, where a system is conceived of as a set of interdependent, interacting components that form a complex, functionally defined whole.

The expression ‘systemic risk’ is the conjunction of two terms. I have advocated above a specific understanding of ‘risk’ as a multi-dimensional concept. The adjective ‘systemic’ derives from the noun ‘system’, which may be taken to refer to a set of interdependent, interacting components that form a complex, functionally defined whole. A lake’s ecosystem, a national electricity supply grid and the global financial markets are examples of systems in this sense.<sup>11</sup> A systemic risk may then, most generally, be understood as a risk that relates to a system. Any definition of systemic risk has to be compatible with this general understanding in order to meet the criterion of semantic adequacy.

But this basic characterisation is ambiguous. The expression ‘relating to’ could mean ‘threatening damage to’ or ‘originating in’. In fact, both senses are relevant, but need to be elaborated upon. In responding to Linarelli, I have argued that some restriction has to be placed on the origin of a risk for it to qualify as systemic. But requiring risks to be endogenous to a system, i.e. to arise from within it, would be too restrictive because few systems are fully isolated, so most changes within them are at least partly caused by external events. Instead, I propose that we include risks external to a system among the systemic risks as long as they are part of the ordinary background noise of the given system, where the ordinary background noise comprises all those shocks to the system that are normally taken into consideration in assessing it or in improving its resilience.<sup>12</sup> In applying this criterion, we can exclude risks that are freakish (and therefore deemed too rare to consider), global in effect (and therefore deemed non-mitigable) or largely innocuous (and therefore deemed irrelevant).

In criticising James, I have argued that some restriction has to be placed on the forms of harm threatened by systemic risks. I propose that systemic risks need to be risks that endanger or threaten the functioning of the given system. This way, we can exclude risks that are entirely local (relative to a given system) or merely threaten unwanted events that are external (relative to a given system). This is important because such risks differ in structure. For example, they will typically be much more

predictable because they do not rely on complex feedback-loops and tipping points that are characteristic of systemic risks elsewhere. As James rightly observes, a system can “[...] unravel with little risk of harm to anyone” and “[...] be dangerous in relative stability” (2017, 240). But this does not undermine the account proposed here. A system whose failure would not harm anyone would simply not give rise to relevant systemic risks. Moreover, while a system might be dangerous even while functioning well, this danger would differ in structure. One could argue that the global economic system imposes a risk of falling into poverty on many individuals even while functioning well. While this risk also results from a complex causal interaction, it would not involve the tipping points characteristic of systemic risks. So, while it might share some relevant features, it will not share all of them.

On this basis, I arrive at the definition proposed. Two things should be noted. First, one might think that this definition remains vague, because it is not always clear what belongs to a system. Are the migrating birds that seasonally rest on the lake to be counted as part of the lake’s eco-system? What about their predators who live far away? In some sense, the demarcation of a system will always remain vague. But this neither implies that it will be arbitrary, nor that there will be misunderstandings. We do in fact use the term ‘system’ successfully in many contexts, and what is to be included when considering a system will often be quite clear from the context. Second, and relatedly, systems can be interlaced and can thus be identified at different levels. A lake’s ecosystem, for instance, is part of a regional ecosystem, which is itself part of a global ecosystem. Because risks might be systemic at one level, but not at the other, it is generally relevant to specify the system relative to which they are evaluated.

### *Characteristic Structural Features of Systemic Risks*

It is clear that, given the account of systemic risk introduced, not all systemic risks will be alike. The risk of severe climate change and the risk of

the breakdown of the financial markets would both qualify as systemic risks, but differ greatly, e.g. with regard to the extent of potential damage, the reversibility of the damage and their time-horizon. But these differences can be captured by the different dimensions of risk, distinguished quite generally in the first part of the paper. What is relevant is that while systemic risks will not all be alike, they will share a number of important features that set them apart. Because these features arise from their origin in a system consisting of interdependent components, I shall argue, they cannot be adequately accounted for by the dimensions of risk.

In this section, I characterise what I take to be the three most substantive structural features of systemic risks.<sup>13</sup> First, the interdependence of the components of a system is likely to result in externalities, spill-over-effects and feedback-loops. While the interdependence may prove beneficial by functioning as an internal compensation device, it can also turn into a reinforcement or contagion mechanism, whereby a small fault in a minor component can trigger a cascade of failures leading to a breakdown of the entire system. For an illustration, consider a road network. If an accident causes a traffic jam in one place, drivers can normally reroute, so other roads compensate for the failure of one link. However, if traffic volume is already high on the other roads, the rerouted drivers may cause a spread of congestion to the other roads. In this case, the interdependence reinforces the initial failure.

Second, the existence of far-ranging feedback-mechanisms tends to result in the system having few stable equilibria. Due to stabilisation and reinforcement mechanisms, the outcomes are neither continuously, nor normally distributed, but erratic. As a result, we need to expect something akin to phase-transitions, where a system tips at some point and then moves, perhaps rapidly, to another, possibly distant equilibrium. Again, the road network illustrates this. Traffic running smoothly is one stable equilibrium; under normal circumstances, if congestion blocks cars on one road, drivers can divert, so other roads compensate for the missing link and the system returns to its initial equilibrium. Widespread traffic-jam is

another stable equilibrium: If traffic starts moving smoothly on one road, other drivers divert to it, reintroducing congestion, so the system returns to its initial equilibrium. In between, there might be other equilibria, but they will inherently unstable. If traffic is merely slow, we expect it to return to normal speed when the number of vehicles decreases or to result in congestion if it increases – so, in the case of a sufficiently large shock, we expect the network to revert to one of the stable equilibria.

Third, systemic risks give rise to strategic interaction. In the road network, what is the best route for a driver depends on the decisions other drivers take, as it is *ceteris paribus* better for each driver to be on an emptier road. So, a good strategy will make his decisions conditional on his expectations about the behaviour of others. No such strategic behaviour is present in simpler cases of risk. Whether or not to place a bet at a roulette table is independent of whether other people place bets. As a result, systemic risks give rise to a particular coordination problem that is akin to the tragedy of the commons. Typically, everyone could benefit from a regulation of behaviour, although each agent would also have an incentive to deviate from it.

### *The Distinctive Normative Considerations to which Systemic Risks Give Rise*

The characteristic structural features shared by systemic risks give rise to a number of distinctive normative considerations. In evaluating them, it is important to distinguish three types of systems. There are systems in which human agents are important components (e.g. the global financial system). There are systems on which humanity has an influence although human agents are not important components (e.g. the ecosystem of a remote, uninhabited island). There are systems whose functioning is beyond the control of humans (the solar system). The third type may here be ignored as it does not give rise to normative considerations. The second type does not give rise to concerns about the interaction among human agents and does not give rise to normative considerations that are not also present in

instances of non-systemic risks. So, I will focus here on the first type of systemic risk. In asserting that systemic risks give rise to four distinctive normative considerations, I do not claim that such considerations never arise elsewhere, but that they do not typically arise in this combination.

First, in the context of systemic risks, the behaviour of each individual agent typically has only a negligible impact on the overall level of risk. This gives rise to the question to what extent his or her individual behaviour can be deemed problematic and to what extent he or she can be deemed blameworthy. Whether or not a typical citizen took out a risky subprime mortgage prior to 2007 did not perceptibly affect the level of the risk of a global financial crisis; but the joint behaviour of numerous citizens who took out such loans did. As James observes, we face a peculiar situation: “Even if a system creates objectionable risks [...], it seems that each agent and action within the system might nevertheless be morally innocent” (2017, 241). Systemic risks then give rise to a particular normative question. To what extent can one be morally required to refrain from participating in an activity where participating in it is privately beneficial and has only a negligible effect on the level of risk of harm imposed on others, but where the activity nonetheless results in the imposition of a serious risk? For a classical approach to solving such collective action problems, see, for example, Ostrom (1990).

Second, and relatedly, due to the causal complexity, neither the occurrence of a risk, nor its materialisation can typically be causally attributed to any specific agent or small group of agents. This undermines the ability to assign responsibility as well as the possibility of mandating compensation. For instance, only a few people can be assigned particular responsibility for bringing about the financial crisis. This gives rise to the question in what ways and according to which criteria other forms of responsibility can be assigned, for instance the responsibility to bring about a change in the structural pattern that gives rise to the systemic risk. Similar questions also arise in the context of collective behaviour that results in the imposition of harm (rather than

risk) and have been prominently addressed for instance by Kutz (2000) and Young (2011).

Third, when confronting systemic risks, agents typically find themselves in a particularly disadvantaged epistemic position. More specifically, uncertainty about the set of possible outcomes and their respective likelihoods is particularly pronounced because interdependencies as well as the complex strategic interaction among agents might render the behaviour of the system erratic and difficult to predict. As a result, it will be particularly challenging to put reliable estimates on the likelihood of outcomes conditional on certain actions. Evaluating the resilience of the financial system for instance is infamously difficult. This does not only result in a challenge to the risk analyst, but also to the moral philosopher, because it raises the question of how precautionous one should be. The question how to act in the face of severe or deep uncertainty of course also arises and has been addressed in the context of non-systemic risks (see, for example, Lempert, Popper and Bankes 2003).

Fourth, in the case of systemic risk, insurance will often be unavailable or at least imperfectly available. This is true in particular where a risk is global (relative to a confined group of people). The principle of (intra-temporal) insurance is based on the method of pooling risks. Those unaffected by a negative event share their luck with those who suffered. But in the case of systemic risks, outcomes are typically highly correlated. This reduces the extent to which risks can be pooled. The financial crisis caused a global economic depression, simultaneously affecting numerous people (though not to the same extent). In the case of the most dramatic and far-reaching such risks, no insurance is feasible at all. If dramatic global warming renders earth inhabitable, everyone then living is doomed. Generally, in the context of systemic risks emphasis has to be placed on the ex-ante regulation of the risk rather than on the ex-post redistribution of outcomes. In this regard, systemic risks resemble risks to public or common goods, which are also imperfectly insurable. For a discussion of risk in such contexts, see, for example, Shrader-Frechette (1991).

Systemic risks give rise to a number of important normative considerations. Individually, each of these considerations may also arise in other contexts. Moreover, as I have indicated, each of them has been addressed, often in texts that have by now attained the status of classics. What is distinctive about systemic risks however is that they give rise to all these considerations at the same time. This complicates the matter as it requires a combination of these different approaches.

*The Relevance of the Notion ‘Systemic Risk’*

The risk of a collapse of the global financial markets and of global climate change illustrate that some of the most significant risks faced by societies today arise from the complex interaction of a great number of agents. These risks are special. They have non-linearly distributed outcomes, are difficult to estimate, they give rise to strategic interaction, the harm they threaten might not be easily reversible and difficult to insure against. I have here proposed that we use the notion ‘systemic risks’ to refer to such risks, which arise from and endanger the functioning of a system. This term is useful because it allows one to refer to a particular group of risks that give rise to a distinctive combination of normative considerations and to do so in a way that makes sense to the ordinary language user. Since systemic risks vary considerably according to their dimensions, they cannot be jointly referred to by reference to the dimensions of risk alone.

The risks I have here identified as systemic risks have, with the exception of the risk of climate change, so far received rather little attention in philosophy. But systemic risks are of growing relevance. The world we inhabit has, at the global as well as the local level, become increasingly interconnected. As a result, there now exists more of the complex interaction that gives rise to systemic risks. For instance, the more integrated financial markets, the easier it is for a local shock in one market to transmit to other markets. Given the process of globalisation, there is indeed a tendency of systemic risks to become global. But they

need not be global. The risk of a failure of a city's waste-disposal system or of a regional road network might be systemic. The great practical relevance of such risks provides one practical reason for addressing them. The fact that they are of particular theoretical interest provides another. Because systemic risks give rise to a number of distinctive normative concerns, which have so far not been addressed in combination, thinking about systemic risks might give rise novel ideas.

#### IV. CONCLUSION

The present contribution started out from the observation that there is a considerable diversity among risks and that this diversity is too infrequently acknowledged by philosophers. While the new emphasis of moral philosophers on decisions made under conditions of risk was welcomed, it was argued that we need to develop a more sophisticated conceptual framework for adequately addressing diverse risks that give rise to distinctive normative considerations. This article has attempted to contribute to such a development in two ways. First, it argued that 'risk' should be understood as a multi-dimensional concept or, more precisely, as a concept that always has to be evaluated in a qualitative, a quantitative and an epistemic dimension and may in addition be evaluated in further dimensions. Second, the article has proposed a particular understanding of the notion 'systemic risk' to characterise some of the most pressing risks societies face. The two proposals are connected in two ways. The notion 'systemic risk' relies on the account of risk developed in the first part while at the same time complementing it. The structural features of systemic risks, which give rise to particular normative considerations, are not adequately captured by the idea of dimensions of risk. While systemic risks can be evaluated in these dimensions, they form a distinct group of risks that it is important to address. In advancing these two proposals, the paper understands itself as a contribution to the ethics of risk – not by addressing the question how to handle a particular risk, but by contributing to our conceptual tools for doing so.<sup>14</sup>

## WORKS CITED

- Altham, James E. J. 1983. "Ethics of Risk." *Proceedings of the Aristotelian Society* 84: 15-29.
- Aven, Terje. 2015. *Risk Analysis*. Chichester: John Wiley.
- Aven, Terje, and Ortwin Renn. 2009. "On Risk Defined as an Event Where the Outcome is Uncertain." *Journal of Risk Research* 12/1: 1-11.
- Boholm, Max, Niklas Möller, and Sven Ove Hansson. 2016. "The Concepts of Risk, Safety, and Security: Applications in Everyday Language." *Risk Analysis* 36/2: 320-338.
- Bostrom, Nick, and Milan M. Cirkovic. 2011. *Global Catastrophic Risks*. Oxford: Oxford University Press.
- Bradley, Richard, and Mareile Drechsler. 2014. "Types of Uncertainty." *Erkenntnis* 79/6: 1225-1248.
- Cline, William R. 1984. *International Debt: Systemic Risk and Policy Response*. Washington D.C.: Institute for International Economics.
- Foot, Philippa. 1967. "The Problem of Abortion and the Doctrine of the Double Effect." *Oxford Review* 5: 5-15.
- Fouque, Jean-Pierre, and Joseph A. Langsam. 2013. *Handbook on Systemic Risk*. Cambridge: Cambridge University Press.
- Frick, Johann. 2015. "Contractualism and Social Risk." *Philosophy & Public Affairs* 43/3: 175-223.
- Fried, Barbara H. 2012. "What Does Matter? The Case for Killing the Trolley Problem (Or Letting It Die)." *Philosophical Quarterly* 62/248: 505-529.
- Hansson, Sven Ove. 1989. "Dimensions of Risk." *Risk Analysis* 9/1: 107-112.
- Hansson, Sven Ove. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. London: Palgrave Macmillan.
- Hansson, Sven Ove. 2014. "Risk." In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. <https://plato.stanford.edu/entries/risk/> [accessed July 7, 2018].
- Heinemann, Simone. 2014. *Ethik der Finanzmarktrisiken am Beispiel des Finanzderivatehandels*. Münster: Mentis.
- James, Aaron. 2017. "The Distinctive Significance of Systemic Risk." *Ratio Juris* 30/3: 239-258.
- Kaufman, George G. 2000. "Banking and Currency Crisis and Systemic Risk: A Taxonomy and Review." *Financial Markets, Institutions & Instruments* 9/2: 69-131.
- Knight, Frank Hyneman. 1935. *Risk, Uncertainty and Profit*. London: The London School of Economics and Political Science.
- Kutz, Christopher. 2000. *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press.
- Lempert, Robert J., Steven W. Popper, and Steven C. Bankes. 2003. *Shaping the next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*. Santa Monica, CA: Rand Pardee Center.

- Linarelli, John. 2017. “Luck, Justice and Systemic Financial Risk.” *Journal of Applied Philosophy* 34/3: 331-352.
- Organization for Economic Co-operation and Development (OECD). 2003. *Emerging Risks in the 21<sup>st</sup> Century: An Agenda for Action*. Paris. <http://www.oecd.org/futures/globalprospects/emergingrisksinthe21stcenturyanagendaforaction.htm> [accessed July 7, 2018].
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action. The Political Economy of Institutions and Decisions*. Cambridge: Cambridge University Press.
- Renn, Ortwin. 2014. “New Challenges for Risk Analysis: Systemic Risks.” *Japanese Journal of Risk Analysis* 24/3: 157-160.
- Renn, Ortwin, and Florian Keil. 2008. “Systemische Risiken: Versuch Einer Charakterisierung.” *GALA - Ecological Perspectives for Science and Society* 17/4: 349-354.
- Renn, Ortwin, and Andreas Klinke. 2004. “Systemic Risks: A New Challenge for Risk Management.” *EMBO Reports* 5/S1: S41-S46.
- Renn, Ortwin, Klaus Lucas, Armin Haas, and Carlo Jaeger. 2017. “Things Are Different Today: The Challenge of Global Systemic Risks.” *Journal of Risk Research* December: 1-15.
- Rosa, Eugene A. 1998. “Metatheoretical Foundations for Post-Normal Risk.” *Journal of Risk Research* 1/1: 15-44.
- Schellnhuber et al., Hans-Joachim. 1998. “World in Transition: Strategies for Managing Global Environmental Risks. Annual Report 1998.” German Advisory Board for Global Change. [http://www.wbgu.de/fileadmin/user\\_upload/wbgu.de/templates/dateien/veroeffentlichungen/hauptgutachten/jg1998/wbgu\\_jg1998\\_engl.pdf](http://www.wbgu.de/fileadmin/user_upload/wbgu.de/templates/dateien/veroeffentlichungen/hauptgutachten/jg1998/wbgu_jg1998_engl.pdf).
- Schwarcz, Steven L. 2008. “Systemic Risk.” *Georgetown Law Journal* 97/1: 193-249.
- Shrader-Frechette, Kristin S. 1991. *Risk and Rationality: Philosophical Foundations for Populist Reforms*. Berkeley, CA: University of California Press.
- Shue, Henry. 1978. “Torture.” *Philosophy & Public Affairs* 7-2: 124-143.
- Silva, Walmir, Herbert Kimura, and Vinicius Amorim Sobreiro. 2017. “An Analysis of the Literature on Systemic Financial Risk: A Survey.” *Journal of Financial Stability* 28: 91-114.
- Williams, Bernard. 1985. “A Critique of Utilitarianism.” In J.J.C Smart and Bernard Williams. *Utilitarianism For and Against*, 77-150. London: Cambridge University Press.
- Young, Iris Marion. 2011. *Responsibility for Justice*. New York: Oxford University Press.

## NOTES

1. For an early attempt to introduce questions of risk into ethics, also see Altham (1983).
2. One could add further senses, e.g. the sense of risk as a measure of variance, as in the claim that lottery A involves a greater risk than lottery B where A and B have the same expected value, but A’s variance is greater.

3. One could argue that we never, strictly speaking, act under conditions of risk in this sense as something unforeseen may always happen. But even if this notion of risk relies on an idealisation, it might still serve a purpose in permitting a distinction between cases of subjectively perceived uncertainty (throwing dice one believes are loaded) and cases featuring only inescapable background uncertainty (throwing dice one believes are fair).

4. Note that the epistemic dimension also exists on a Bayesian account. It evades the need for a distinction between risk and uncertainty, but the distribution of credences measures the strength of the epistemic position.

5. He states: “The risk concept  $(C,U)$  covers (i) that the activity leads to some consequences  $C$ , and (ii) that these consequences are not known  $(U)$ ” (Aven 2015, 16).

6. Strictly speaking, James only provides a definition of what it is for a risk to be “created *systematically*” (cf. James 2017, 240; italics original); but since the entire article focusses on the topic of systemic risk and he uses the term “systemic risk” in the title, it seems justified to assume that he takes a systemic risk to simply be a risk that has been created systematically.

7. In fact, I think that this question points to a lacuna in the literature. The introduction of technical terms amounts to a form of conceptual engineering, on which little has been written, in particular in practical philosophy.

8. One could argue that it also violates the criterion of external adequacy because in other areas, for instance in ethics, non-global systemic risks give rise to similar (though not identical) normative considerations, so they require a broader notion of systemic risk.

9. In fact, Renn *et al.* use ‘global systemic risks’ in the title, but then talk of systemic risks in general.

10. They claim that paradigmatic examples establish “[...] a web of family resemblances” and reject the idea that “[...] a single standardized definition of systemic risks” can be given (Renn *et al.* 2017, 3).

11. While the term ‘system’ can also refer to other entities, such as an organised set of theoretical principles (‘the Newtonian system’) or a scheme for measuring (‘the metric system’), this ambiguity can here be ignored. Only non-theoretical, non-classificatory systems that rest on the causal interaction of their components are of interest in the context of a discussion of risk.

12. This proposal is similar in spirit, but different in detail to that of Renn *et al.*, who limit consideration of exogenous processes by requiring them to modify the internal dynamics of a system (2017, 3).

13. These overlap with what Renn *et al.* (2017) take to be defining properties.

14. I would like to thank Francis Cheneval and Juri Viehoff for commenting on an earlier version of this paper. Moreover, I would like to thank two anonymous reviewers for *Ethical Perspectives* for very helpful comments. Financial support from the Swiss National Science Foundation (Project Number 162776) is gratefully acknowledged.

# *Risk, Science and Policy: A Treacherous Triangle*

Sven Ove Hansson

**ABSTRACT.** Decisions on risk are often frustrated by conflicts and uncertainties at the science-policy interface. This article shows that there is considerable scope for improving the decision-making process and in particular its use of science. Philosophical tools and distinctions can have a key role in eliminating or coming to grips with risk-related uncertainties. Many of the so-called scientific controversies are in fact not conflicts within science, but between science and science denial or other forms of pseudoscience. In these cases, the alleged uncertainties can be dissolved by throwing out pseudoscientific claims and fabricated controversies. A significant part of the remaining uncertainty is not decision-relevant although it seems to be so. Two simple symmetry tests are introduced for investigating whether an uncertainty is specific enough to be more troublesome for the option it has been associated with than for other options in the decision. Finally, decision theory and ethics provide means for dealing with the remaining decision-relevant scientific uncertainties. One of the ethical tools that can be used for this purpose is hypothetical retrospection. This method consists in systematically deliberating on how we will in the future evaluate the decisions that we make today, under various assumptions about other potential future developments. The purpose of this process is to come as close as possible to the ideal of making a decision that will be ethically defensible whatever happens.

**KEYWORDS.** Uncertainty, fact-finding practices, science-policy interface, pseudoscience, science denial, risk policy

## I. INTRODUCTION

Science has a key role in many controversial policy issues, not least those concerning risks to public health and the environment. Ideally, one would expect the scientific input to provide a common factual ground in issues of risk, so that policy discussions can focus on the underlying values and the desired outcomes of the decision. In this way, science would transform disorderly policy debates into the more structured and constructive type of discussion that we are used to in science. But in

actual practice, the influence often goes in the opposite direction. When risk-related science comes in close contact with policy-making in controversial areas, it tends to become heavily affected by the policy controversies (Wagner 1995). Disagreements in scientific issues flare up, uncertainties are inflated, and scientific discussions become infected with divisive partisan viewpoints. Again and again, risk, science, and policy come across as an inharmonious trio, whose joint appearance is commonly accompanied by unmanageable uncertainties and entrenched social conflicts.

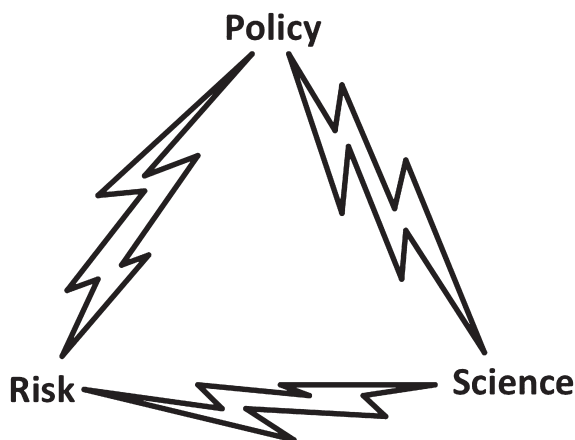


Figure 1. The treacherous triangle.

This article will make two major points. First, the treacherous triangle risk—science—policy need not be as hopelessly conflict-ridden and uncertainty-laden as it often seems to be. Secondly, philosophical tools and distinctions can have a key role in improving the situation. This includes tools from at least three areas of philosophy: philosophy of science, decision theory, and ethics. Section II provides the necessary background on the nature of science and scientific knowledge. Section III introduces a classification of uncertainties and controversies. Section IV shows that much of the apparent uncertainty in risk policy issues can be dissolved by throwing out pseudoscientific claims and fabricated controversies. Section

V makes it clear that much of the remaining uncertainty is not decision-relevant, and Section VI exemplifies that the decision-relevant uncertainty can be managed with the help of appropriate decision-theoretical and ethical tools. Section VII concludes.

## II. THE NON-UNIQUENESS OF SCIENTIFIC KNOWLEDGE

Science is often depicted as based on unique and superior ways of thinking that diverge radically from how we think in everyday life. Sandra Harding refers to this as “the belief that Western sciences alone among all human knowledge systems are capable of grasping reality in its own terms—‘cutting nature at its joints,’ as philosophers of science typically enjoy referring to the matter” (2008, 3-4). The prevalence of this approach can be seen in common textbooks in the philosophy of science, where scientific knowledge is treated as an entity separate from other forms of knowledge. This also appears to be the most common viewpoint in the research literature. However, it is clearly a misrepresentation of the scientific enterprise. Fundamentally, science is based on the same two sources of knowledge that we use in our everyday lives, namely our senses and our reasoning abilities.<sup>1</sup> What makes science special is not its basic means of acquiring knowledge but the specialized and persevering ways in which they are used.

Furthermore, science is not the only human activity that is devoted to refining and applying the means to knowledge that are available to us. Science is one of what I have proposed to call the *fact-finding practices* of human societies (Hansson 2018a; 2018b). A fact-finding practice is a human endeavour that has a strong focus on obtaining joint and reliable knowledge on some subject-matter. The deliberations of a jury or a panel of judges before rendering a guilty/not guilty verdict is another important example, and so are the activities of investigative journalists, car mechanics trying to find out what is wrong with an engine, auditors performing a financial audit, etc.<sup>2</sup> In addition to the many professionalized fact-finding practices, we also have many such practices in our everyday lives.

For instance, two people looking for a lost object, such as a mislaid bunch of keys, exhibit many of the behaviours that are also characteristic of professional fact-finding practices. They set up hypotheses (“Maybe you left the keys on the front seat?”), test those hypotheses, draw conclusions from their observations etc. (Hansson 2018a).

Fact-finding practices can be found in all human societies and cultures, certainly not excluding illiterate ones. For instance, anthropologists have recorded surprisingly science-like conversations among hunters in the Kalahari desert who follow a prey for many hours in “a process of creative problem solving in which hypotheses are continually tested against spoor evidence, rejecting those which do not stand up and replacing them with better hypotheses” (Liebenberg 1990, 71; cf. Liebenberg 2013).

Fact-finding practices are characterized by endeavours to separate facts and values from each other as far as possible. When fact-finding is applied to contentious topics, this separation can be difficult and even impossible to fully achieve. This applies to all fact-finding practices (not only science). For instance, the difficulties of separating facts from values in treatments of economic policy are much the same in economic journalism and economic research. The best way to deal with fact-value problems is also essentially the same in the different fact-finding practices: separate out the facts as far as possible, and report the remaining intermixture as accurately as possible. For instance, in countries with well-functioning labour market statistics it is possible (both in economic research and economic journalism) to provide a reasonably value-independent answer to the question “How many persons in the ages 15-24 do not have at least a half-time time job and are not registered for at least half-time time studies?”. However, if there is a political conflict over the ‘correct’ definition of youth unemployment, then it may still be impossible to answer the question “How many youths are unemployed?” in a reasonably value-independent way. It is essential to make the distinction between these two cases as clearly as possible, and also to make sure that the value-dependence of the latter question does not spill over to the

former. These distinctions are often not made as they should. As noted by Doppelt, “[t]he claim that science, as we know it, is not and cannot be value free” has gained “an almost theological status for many scholars today” (2007, 189), but it severely blurs the distinctions between the cases when science can respectively cannot provide sufficiently value-independent information. There is a need for systematic work to improve both the fact-value separation and the way it is reported. (Hansson 1995; 2017b; Douglas 2009)

But in spite of the important similarities among the different fact-finding practices, there are of course noteworthy differences between science and the others. Perhaps most important among these are its high degree of specialization, its strivings for comprehensiveness, its resourcefulness, and (today) its global network. Importantly, the term ‘Western science’ for today’s global science is a Eurocentric misnomer. Science as we know it would not have been possible without the contributions of ancient Iraqi (Mesopotamian) astronomers, Indian mathematicians, Arabic physicists, and indigenous people all around the world who provided European explorers with information about the fauna and flora of their homelands.

Three important features of science come out clearly once we have understood its origin and its connections with the wider category of fact-finding practices. First, like the other fact-finding practices referred to above, *science is a collective enterprise in which ideally – i.e. at its best – the strength of an argument is all that counts, and consequently, the social status of the person who puts it forward counts for nothing*. This feature of science is essential for its credibility in the eyes of the public. Unfortunately, the public sees very little of the non-authoritarian and open-minded daily discussions among scientists.

Secondly, like the other fact-finding practices, science *aims at finding out the actual state of things, which must be the same for all of us*. Hunters in the Kalahari desert take this for granted; they would not accept the standpoint that “the antelope we have been chasing is behind that hill to our left in your world, but in my world it is far away in the opposite direction”.

If you and I are looking for your lost keys, you would probably not find it very helpful if I told you that “in my world, the keys are on the kitchen table, but in your world you have lost them on the bus”. Few of us would be satisfied if a police investigation of a suspected sexual assault concluded that “it is the victim’s truth that this was a rape, but it is also the defendant’s truth that the intercourse was consensual”. In all these cases, we require that the fact-finding practice should aim at *the* truth, not just *someone’s* ‘truth’. Obviously, the same applies to science. To mention just one example, we expect climatologists to tell us, as accurately as they can, the truth about climate change. They would not be doing their job if they told us that there are several, equally valid but widely different truths in the matter, for us to choose among.

The single-truth assumption, as we may call it, seems to be so self-evident in all other fact-finding practices than science that it is not even talked about. However, its application in science has repeatedly been questioned. Relativism about scientific knowledge has become a popular stance in some allegedly intellectual circles, not least in some segments of social science. There can be little doubt that the relativist attacks on science have facilitated the current denialist onslaught on climate and environmental science (Latour 2004; Banning 2009; Lynch 2016; Collins *et al.* 2017; Fuller 2017). This intellectual collapse would probably not have been possible if the close connections between science and other fact-finding practices had been generally known and acknowledged. Once the standard arguments for relativism about scientific knowledge is transferred for instance to investigative journalism, financial auditing, or animal tracking, it will immediately be seen how easily relativism can be refuted.

Thirdly, like other fact-finding practices, science (when it works properly) is *open to corrections and improvements of its knowledge claims*. A scientific statement can never achieve the incontrovertible status that is often ascribed to religious dogmas. The highest status that a scientific statement can achieve is that of *provisional knowledge*. This means essentially that we currently have no reason to doubt the statement in question, and therefore

we treat it as a true statement. However, if reasons to doubt it (unexpectedly) come to light, then we must be prepared to change our minds.

Again, a comparison with other fact-finding practices will show how closely scientific knowledge connects to other types of knowledge. If there is no reasonable doubt that the defendant committed the crime, then the judicial system will treat him (her) as guilty. However, many legal systems have provisions for a trial *de novo* when new facts are discovered that could lead to a different verdict (Oldfather 2008). Similarly, if journalists obtain evidence proving beyond reasonable doubt that a politician has engaged in corrupt practices, they will publish. If surprising new information acquits the politician, media ethics requires the publication of corrections and apologies.

Most of what we consider to be our everyday knowledge is provisional in the same sense: we take something to be true if we currently have no reason to doubt it, but (surprising) new information can make us change our minds. For instance, when I left home this morning I considered myself to know the shortest way to the local Metro station. However, it was perfectly possible that a few minutes later I would have found myself to be wrong since my passage was blocked for instance by a sink-hole, the shooting of a film, or a criminal investigation. Just like scientific knowledge, everyday knowledge is certain in the sense of currently not being doubted, but neither of them is certain in the sense of being immune to confuting evidence. This applies to all forms of rationally defensible empirical knowledge. There are no grounds whatsoever for sorting out scientific knowledge as particularly untrustworthy because of this feature (as is often done with remarks such as “scientists have been wrong before”, etc.).<sup>3</sup>

### III. A CLASSIFICATION OF UNCERTAINTIES

We can classify the scientific uncertainties relating to a decision according to two independent distinctions. One is the distinction between conflict-based

and consensus-based uncertainties. Scientific uncertainty on an issue is conflict-based if there is genuine scientific controversy over the issue in question.<sup>4</sup> The uncertainty is consensus-based if there is scientific consensus that science cannot currently determine which of the alternative standpoints is correct. The difference is important since we have considerable experience showing that (what appear to be) conflict-based scientific uncertainties are more difficult to deal with than those that experts agree upon. A scientific consensus provides a common ground for policy discussions, even if it is (part of) the consensus that some answers are not available.

|                 | Decision-relevant                                                                 | Decision-irrelevant                                                               |
|-----------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| Conflict-based  |  |  |
| Consensus-based |  |  |

|                                                                                     |                                                              |
|-------------------------------------------------------------------------------------|--------------------------------------------------------------|
|   | Highly problematic, often causes a stalemate                 |
|  | Makes the decision difficult, but is usually manageable      |
|  | Causes no problems if the decision-irrelevance is recognized |

Figure 2. A classification of uncertainties, based on their decision relevance and whether they are conflict-based or consensus-based.

The other distinction is that between decision-relevant and decision-irrelevant uncertainty. An uncertainty is decision-relevant if a resolution of the uncertainty can have an impact on the decision. For instance, suppose a decision has to be made on whether or not a particular pesticide should be permitted. If it is uncertain whether it is toxic to pollinating insects, then that is a decision-relevant uncertainty. On the other hand, if there is consensus that the pesticide has such effects, but uncertainty

about the biochemical mechanisms, then the latter uncertainty is not (at least not *per se*) decision-relevant for this decision.

The two distinctions can be combined into a four-fold classification scheme, as shown in Figure 2. The most problematic uncertainties are those that are both conflict-based and decision-relevant. In Section IV we will see that such uncertainties are much less common than they seem to be. The other problematic group consists of the consensus-based, decision-relevant uncertainties. In Section V we will see that they too are less common than they are usually assumed to be.

#### IV. SCIENTIFIC CONTROVERSIES ARE NOT AS COMMON AS THEY SEEM

The purpose of this section is to show that several societal mechanisms combine to engender gross exaggerations of the prevalence of scientific controversies in risk-related policy issues.<sup>5</sup>

*First thesis: Genuine scientific controversies in risk issues are much less common than they appear to be. This will become clear once we throw out the pseudoscience from policy discussions.*

A pseudoscience is a practice that satisfies the following criteria:

- i. It pertains to an issue within the domains of science in the broad sense (the criterion of scientific domain).<sup>6</sup>
- ii. It suffers from such a severe lack of reliability that it cannot be trusted at all (the criterion of unreliability).
- iii. It is part of a doctrine whose major proponents try to create the impression that it represents the most reliable knowledge on its subject matter (the criterion of deviant doctrine) (Hansson 2013b, 70-71).

The doctrinal component in the third part of the definition distinguishes between pseudoscience and other violations of the requirements of (good)

science. Mistakes due to equipment failure, human failure in honest attempts to produce reliable results, and most forms of scientific fraud are not commonly called pseudoscience. Although such aberrations can have severe negative effects on the scientific process, once exposed they are usually much easier to dispose of than pseudoscience since they lack ideologically motivated defenders.

There are two major forms of pseudoscience, distinguished by the character of the doctrinal component: pseudoscientific theory promotion and science denial(ism) (Hansson 2017a). Pseudoscientific theory promotion consists in the promulgation of some specific theory or idea that is in conflict with science. Some examples are astrology, biorhythms, crystal healing, graphology, homeopathy, Lysenkoism, scientology, and ufology. The espousal of these teachings usually involves dissent from some domains of science, but that is not their primary objective.

In contrast, science denial has as its supreme goal to overthrow some specific scientific theory or area of scientific inquiry. Typical examples are holocaust denialism, relativity theory denialism, AIDS denialism, vaccination denialism, tobacco disease denialism, germ theory denialism, and climate change denialism. Creationism is also a form of science denialism, since it is a religiously motivated attack on the scientific study of evolution, rather than an attempt to advance any theory about biological subject matter. Although pseudoscientific theory promotion and science denial are far from sharply delimited from each other, most forms of pseudoscience can be assigned fairly unequivocally to one of these two categories.

Both forms of pseudoscience have a prominent presence in policy discussions concerning risks to public health and the environment. In particular, various forms of science denialism have been propagated by powerful economic interests, most notably tobacco science denialism by the tobacco industry and climate science denial by the fossil fuel industry (Michaels 2008; Oreskes and Conway 2010). It is important to recognize the major strategies used by these and other promoters of pseudoscientific claims in risk issues.

The *fabrication of false scientific controversies* is perhaps the most important of these strategies. It is a method used almost exclusively by science deniers, not by pseudoscientific promoters. Those who primarily advance a specific pseudoscientific theory, such as homeopathy or ancient astronauts, typically downplay their conflicts with established science. Thus, homeopaths and other quacks misleadingly describe their ineffective treatments as ‘complementary’, and ancient astronaut theorists and other pseudo-historians zealously allege support for their claims in legitimate historical scholarship. In contrast, since the aim of science denialism is to defeat certain (sub)areas of science, conflicts with those parts of science are unavoidable. Science denialists usually frame this conflict as a dispute between two scientific standpoints, one of which is their own. In actual fact, it is a conflict between the scientific consensus and their own brand of pseudoscience. The strategy of creating false scientific controversies was already prominent in the denial of relativity theory in the late 1920s (Wazeck 2009, 268-269). It was developed to a high degree of sophistication in the 1950s in the deliberately death-bringing activities of the tobacco companies (Brandt 2007; 2012). It has been taken over by industry-funded deniers of various scientifically established risks to human health and the environment (Michaels 2008; Rudén and Hansson 2008; Oreskes and Conway 2010). Climate science deniers, in particular, have been remarkably successful in making journalists present pseudoscientists like themselves side by side with legitimate climate scientists, treating them as representatives of equally valid scientific standpoints (Boykoff and Boykoff 2004; Boykoff 2008; McKewon 2012).<sup>7</sup>

A distinct but closely related strategy is the *delusive transformation of value conflicts into scientific controversies*. This approach can be illustrated with conflicts on the regulation of chemical substances. For most substances, the available evidence is limited in both scope and quality, leaving considerable uncertainty both about the nature of potential adverse effects and about dose-response relationships. In most cases, if a group of independent toxicologists is tasked with describing the current state of knowledge about

a substance, they will be able to agree. Usually their consensus will describe some effects as established, and others as suspected to various degrees, based on relevant but insufficient evidence. Obviously, widely different policy conclusions can be drawn from such a report, depending on one's value-based standpoints on risk-taking and on the priorities between economic and environmental or health-related considerations. However, there are at least two common mechanisms that blur the distinction between value-based and factual considerations in the assessment of chemical risks.

One of these mechanisms is the organizational integration of science and policy. Ideally, one would expect a committee of scientists to produce a report on the risks associated with an exposure (including the uncertainties), leaving the value-based decisions on regulations and other practical measures to policy-makers. However, policy-makers often prefer to delegate such decisions to the experts. Consequently, committees of scientists are entrusted with the combined task of assessing the toxicity of a substance and determining what exposures are acceptable (Clausen Mork and Hansson 2007). Needless to say, this will make it more difficult for the committee to keep facts and values apart, and consequently increase the risk that their evaluations of the scientific literature will be unduly influenced by their policy standpoints.

The other mechanism comes into play when scientific risk issues become the subject of debates between interest groups. As shown by Wendy Wagner, interest groups often prefer to describe their viewpoints as based on (incontrovertible) science rather than (obviously disputable) values. They “resort to ‘counter-science’ arguments when faced with the alternative course of engaging in unwinnable debates over policy, such as the proper level of risk aversion for protection of public health or the appropriate social cost, expressed in dollars, of an additional cancer” (Wagner 1995, 1658). This will lead to a ‘science charade’ in which parties who disagree thoroughly on the policy issues all present their dispute as a controversy about science. In this way, the public gains the impression that there is much more genuine scientific disagreement in the issue than is actually the case.

Both these mechanisms – organizational integration and the conflict-driven science charade – can operate without any of the parties taking recourse to outright pseudoscience. In cases of substantial scientific uncertainty, a wide range of interpretations are compatible with the available data. If the combatants settle for different interpretations within that range, then they do not make blatantly pseudoscientific claims about the risk, but their combined choices nevertheless create an exaggerated picture of scientific disagreement.<sup>8</sup> However, especially in the latter case (conflict-driven science charades), pseudoscientific claims are far from uncommon.

One of the major argumentation methods used in a wide range of pseudosciences is *cherry-picking*, i.e. the selection of small parts of the scientific evidence that, taken in isolation, seem to support some deviant claim, although the evidence as a whole points conclusively in the other direction. In the 1990s, for instance, the tobacco industry cherry-picked studies of passive smoking with too low statistical power to discover the increased prevalence of deadly diseases among people exposed to passive smoking (Barnes and Bero 1998; Francis, Shea and Samet 2006). Ancient astronaut theorists select small pieces of historical and archaeological information that seem to confirm their claims unless one has access to adequate contextual information (Stiebing 1984; Hein 1981). Similarly, creationists pick out apparently inexplicable paleontological findings without mentioning their received scientific explanation (Gastaldo, Stevanović-Walls and Ware 2004; Matson 2002; Schaferman 1983). Climate science denialists employ the same tactics, for instance by using anomalous local temperature trends that differ from the global trends to ‘disprove’ the reality of global climate change (Tucker 2012, 846-847). The promoters of various pseudoscientific remedies make much use of poorly documented reports about cherry-picked patients. From such reports it can often be concluded that a particular person had a positive health outcome after receiving the treatment in question. However, it cannot be concluded that such positive outcomes are more common among patients who have received the treatment than among those who have not received

it (not to speak of causal connections between treatment and outcomes). As these examples show, the use of cherry-picked data as spurious evidence for scientifically untenable positions contributes much to the construction of bogus conflicts where there is in fact scientific consensus.

A closely related method is that of *dismissing refuting information*. As indicated in Section I, willingness to revise or give up a standpoint when the evidence so requires is an essential characteristic of the scientific ethos, and it is necessary for scientific progress. The pseudosciences lack this mechanism, and their belief systems are therefore more stable than that of science. Stability in this case equals incorrigibility. A comparison between homeopathy and scientific pharmacology can illustrate the difference. Homeopathy was devised in the 1790s and has not changed significantly since then. It was thoroughly disproved through a series of chemical discoveries in the 19<sup>th</sup> century. Most importantly, these discoveries made it abundantly clear that after the long series of dilutions required in the production of homeopathic drugs, there is literally nothing left of the chemical substance supposed to produce a positive health outcome. This is as clear a case of ‘the emperor’s new clothes’ that there ever was. But nevertheless, present-day homeopaths persist in prescribing ‘drugs’ that are so diluted that they are chemically indistinguishable from common water. In contrast, in the same period since the 1790s, scientific pharmacology has gone through several radical changes in its methodology, including the introduction of quantitative chemical analysis, the concept of an active ingredient, the notion of a dose-response relationship, and the use of blinded or double-blinded clinical trials to determine the effects of pharmacological treatments.

Examples of such incorrigibility likewise abound in the other pseudosciences. For instance, as noted by Stephen Weart, in climate science denialism “[t]he old claims that the carbon dioxide spectra was saturated, that urban heat explained the observed global warming, and a hundred others, long since refuted, kept convincing novice readers that claims of global warming were nonsense” (2011, 46). Defending standpoints that

are known to be scientifically untenable is another way in which pseudo-scientists create the illusion of scientific conflict where there is none.

In summary, one of the most pernicious activities of science denialists is the fabrication of false scientific conflicts. Due to the massive denialist campaigns instigated by business interests, such conflicts have a large presence in public discussions on risk issues. Therefore, the scientific basis appears to be much more controversial in risk-related policy-making than in most other policy areas. However, this is not due to the nature of the scientific data but to the unusually strong influence of pseudoscience in this area. By exposing pseudoscience and defending science, we can dissolve the illusory conflicts and create a more extensive common ground for the value-based discussions that are needed in this as well as other policy areas.

## V. MUCH OF THE REMAINING UNCERTAINTY IS NOT DECISION-RELEVANT

*Second thesis: Some of the uncertainties in risk issues that are presented as decision-relevant are in fact not relevant for the decision. This can be shown with two simple symmetry tests, the test of alternative causes and the test of opposite effects.*

The decision-irrelevance of many uncertainties is obvious and unlikely to be disputed. Suppose, for instance, that there is reliable scientific evidence showing that a chemical is a potent carcinogen, but it is uncertain whether or not it is also mildly irritating to the eyes and the respiratory system. This uncertainty will not be decision-relevant for a policy-maker deciding how the use of this substance should be regulated. A regulation that provides adequate protection against the carcinogenic effects is sure to also prevent exposures that would give rise to other, less serious effects of the substance. However, there are also uncertainties that are presented as decision-relevant, and may at first glance seem to be so, but are in fact irrelevant for the decision. Current debates on vaccination against childhood diseases provide a clear example of this.<sup>9</sup>

A decision to vaccinate against a disease will always have to be based on a careful comparison of the risks of vaccination with those of non-vaccination.<sup>10</sup> If we vaccinate, then the persons who receive the vaccine will be exposed to whatever risks there are of harm caused by the vaccine. On the other hand, if we do not vaccinate, then we have to deal with the risk that unvaccinated persons might contract a disease that vaccination would have protected against. (For diseases with human-to-human transmission this risk will also affect other unvaccinated persons than those concerned by the decision.) The common vaccine against childhood diseases, the MMR (measles, mumps, and rubella) vaccine, has been scientifically evaluated in major studies based on comparisons between vaccinated and unvaccinated children (Bloch *et al.* 1985; Maglione *et al.* 2014; Di Pasquale 2016). These studies leave no reasonable doubt that the risks of MMR vaccination are much smaller than those of refraining from the vaccine. However, anti-vaccination activists claim that this and other vaccines have much more serious side effects than has been shown by science (Betsch and Sachse 2013; Kata 2010). Their most prominent claim is the supposition that MMR vaccine causes autism. This claim was put forward by Andrew Wakefield in 1998 in an article that attracted much attention. It became clear very soon after its publication that it was scientifically problematic. Now we know that it was a complete fake. Wakefield had no empirical data at all indicating any connection between the vaccine and autism (Deer 2011a; Deer 2011b; Harrison 2013). In spite of this, however, anti-vaccination activists have continued to claim that there is a connection, or at least remaining scientific uncertainty in the issue (Hobson-West 2007; Dixon and Clarke 2013). In arguing for this claim, they essentially assert that since scientists cannot prove with absolute certainty that there is no connection, the possibility of such a connection must be treated as a real risk.

Obviously, science cannot prove with absolute certainty that this vaccine will never, in any person, causally contribute to autism. No empirical claim can be absolutely certain in the same sense as it is certain that  $2+2$  equals 4. However, even if the uncertainty about this (or any other)

vaccine effect cannot be completely dispelled, the uncertainty-based argument against vaccination can be fully refuted. This is because of two other things that science *can* do. To begin with, science can find out whether autism is more common among vaccinated persons than among comparable unvaccinated ones. Large and well-conducted studies of the alleged connection have shown no difference (in either direction) between vaccinated and unvaccinated individuals. For obvious statistical reasons, these studies do not exclude the possibility of a minuscule connection (in either direction), but practically speaking they demonstrate that if there is a connection it has to be very small (Doja and Roberts 2006; Maglione *et al.* 2014).

Furthermore, science can tell us whether there are any valid reasons to suspect that the vaccine, rather than anything else that happens in a young person's life, can lead to autism. The answer is equally clear: there is no credible reason, such as a plausible mechanism, why the vaccine should have this effect. Again, science does not provide absolute certainty. There might be some potential mechanism that we are not aware of. However, this is not specific for the vaccine. Exactly the same can be said about any other of the many factors in a young person's life, such as riding the merry-go-round, playing with a skipping rope, or eating ice cream with strawberries. Scientifically speaking, each of these activities is just as strong a candidate as the MMR vaccination for being a causal factor contributing to autism. (Arguably, they are in principle stronger candidates, due to lack of epidemiological evidence of the type that is available for the MMR vaccine.) And most importantly: in the sense in which we cannot prove with certainty that the vaccine does not contribute to autism, we cannot refute the alternative supposition that the vaccine contributes to protect against autism. Therefore, even if science has not provided us with absolute certainty that there is no vaccine-autism connection, it has provided us with sufficient evidence to remove that claim from all deliberations on (private or public) decisions on vaccination. In other words, the uncertainty in question is not decision-relevant.

The argument that demonstrated the decision-irrelevance of these uncertainties was based on comparisons between ‘opposite’ uncertainties that are equally supported by the available evidence (but not necessarily communicated with equal fervour). This type of comparisons has been systematized in the form of two symmetry tests for the decision-relevance of uncertainties. (These tests were first proposed in Hansson 2004.)

*The test of alternative causes* consists in investigating whether the uncertainty in question can be defeated by showing that we have at least as strong reasons to consider the possibility that the same effect (or some other effect that is at least as undesirable) will come about if the action under consideration is not performed. In the vaccine case, we applied this test when concluding that we have equally strong reasons for believing that non-vaccination increases the risk of autism as we have for believing that vaccination does so.<sup>11</sup> (‘Equally strong reasons’ here means ‘in both cases essentially no reasons’, but that does not affect the validity of the argument.) The presence of a symmetrical argument that provides equally strong support of the opposite standpoint makes the initially observed uncertainty decision-irrelevant. This is a common situation for highly speculative uncertainties. For instance, the media abound with claims that either increased or decreased intake of various foodstuffs has positive health effects. Many of these assertions have no credible justification whatsoever. In such cases, the opposite claim might just as well be right (Kuske 1983; Lessof 1992, 18-34; Guil-Guerrero 2014). We are therefore well advised to treat these recommendations as decision-irrelevant, i.e. not let them influence what we eat.

*The test of opposite effects* consists in investigating whether an uncertainty concerning a potential risk can be outweighed by some potential benefit that has equal or larger moral weight than the potential risk.<sup>12</sup> Above, we applied this test when we compared the potential negative effects of the MMR vaccine with its much more substantial and thoroughly verified positive effects. The test of opposite effects is particularly helpful in discussions on technologies whose potential positive effects are less well

known than their potential negative effects. Agricultural biotechnology is a case in point (Hansson 2014).

The following case exemplifies that there are also asymmetric uncertainties that neither of the two symmetry tests will nullify.

A company applies for an emission permit to discharge its chemical waste into an adjacent, previously unpolluted lake. The waste in question has no known ecotoxic effects. A local environmental group opposes the application, claiming that the substance may have unknown deleterious effects on organisms in the lake.

In this case, since life in the lake has been evolutionarily adapted to an environment without the waste, it is highly implausible that discharging the waste would have any positive effects on life in the lake. Therefore, neither of the two symmetry tests will nullify the original uncertainty argument.

## VI. WE HAVE MEANS FOR DEALING WITH THE REMAINING UNCERTAINTIES

We can conclude from Sections IV and V that conflict-based and decision-relevant scientific uncertainties are less common in policy issues concerning risk than currently ongoing public debates would make us believe. The purpose of this section is to show that we have means for dealing with the remaining uncertainties. Traditional risk analysis requires probabilistic information and is therefore not suitable to deal with uncertainties that cannot be reduced to probabilities. However, other methods are available.

*Third thesis: Decision-relevant uncertainties in risk issues can be successfully managed with decision-theoretical and ethical tools that do not require probabilistic information.*

In what follows I will present one such tool, hypothetical introspection.<sup>13</sup> For other such tools the reader is referred to (Hansson and Hirsch Hadorn 2016) or the brief summary in (Hansson and Hirsch Hadorn in press).

To be a moral agent means to go beyond the simple ‘me now’ perspective. That is how we morally educate our children; we train them to see things from others’ perspectives and to plan and save for the future. Moral philosophers have devoted considerable efforts to developing and advocating the first of these two extensions of the ‘me now’ perspective, namely seeing one’s own actions from other persons’ perspectives. A large number of argumentation patterns have been constructed to facilitate impartial moral analysis. For decision-making under uncertainty, the extension from ‘now’ to future viewpoints is equally important. The method of hypothetical retrospection has been constructed to perform this extension in a systematized way.

In our everyday lives we often use a simple type of future-directed argument that can be called the ‘foresight argument’. It consists in an attempt to see things the way we will see them at some later point in time. Its simplest applications refer to situations we treat as deterministic. For instance, some of the consequences of drinking excessively tonight can, for practical purposes, be regarded as foreseeable. Thinking in advance about these consequences may well deter a person from drunkenness. In everyday moral reasoning, this way of thinking is often combined with attempts to see things from the perspectives of others. For instance, considering the future consequences for one’s children may well prevent a person from indulging in immediate pleasures with negative long-time consequences.

When the foresight argument is applied to cases with risk or uncertainty, the evaluation of an action will have to take into account more than one future development that can unroll after the action. Since we often tend to neglect some of the potential future consequences of our actions, we need to make a systematic analysis of the various options and their potential consequences, taking care not to omit the negative consequences of convenient or otherwise appealing options. Two examples:

Due to building renovation, a manufacturing department working with inflammable substances has to be temporarily moved to another building for six weeks. Considering the costs of an automatic fire

alarm, the factory manager decides not to install one in the temporary workshop. She would probably not have made that decision if she had carefully considered its worst potential consequence, namely a lethal fire that could have been prevented by a fire alarm.

Currently, spent nuclear fuel from the Swedish nuclear industry is kept in an interim storage facility where it is continuously monitored. A decision has been made to transfer it to a much safer, deep underground final storage facility. Detailed plans for that storage are now under expert and judicial review, after which the government will decide whether to approve them or not. In the latter case, the experts that designed the storage facility will be sent back to the drawing board, and the transfer from interim to final storage will be considerably delayed. At present, there is much discussion about the worst-case scenario of approving the plans, which seems to be corrosion in the copper canisters, leading to radioactive leakage into the groundwater at some future point in time, before the nuclear waste has ceased to be highly toxic. However, the worst-case scenario of the other alternative, not approving the plans, does not even seem to have been identified. One strong candidate is the spread of radioactive material due to an attack or other unusual event in the interim storage facility, taking place when the waste would have been at a safer place if the original plan had been approved. Another strong candidate is an unrelated economic, military, or environmental disaster during the extended preparation time, making the nation incapable of completing the construction of the final storage facility. (Consideration of these scenarios may well have no impact on the decision, but that can only be determined if they are included in the deliberations.)

When preparing an important decision we need to think through each of the major options that we have, and consider what outcomes it can lead

to. In cases of risk or uncertainty, each decision option can be followed by different ‘branches’ of future development. We need to consider each of these branches and think through how we will, in that branch, look back at the decision that we (from this perspective of hypothetical retrospection) made today. For instance, the factory manager in our example should carefully consider how she will react in a hypothetical future situation with fatalities that a fire alarm would have prevented. If we can identify potential future situations in which we would be unable to defend our intended course of action, then that is a good reason to reconsider our plans.

From a moral point of view, we wish to ensure that whichever branch materializes, a posterior evaluation should not lead to the conclusion that what we did was morally wrong. Our decisions should be morally acceptable (permissible) even if things do not go our way. This can also be expressed as a criterion of decision-stability or evaluation-stability: our conviction that the decision was right should not be perturbed by information that reaches us after the decision. In order to achieve this, we have to consider, for each option in a decision, the major future developments that can follow if we choose that option. In other words, we have to perform a systematic hypothetical retrospection.

Ideally, decision-making guided by hypothetical retrospection should always result in a decision that will be acceptable whatever happens. There may be cases in which this cannot be achieved, i.e. cases in which there is no decision alternative that appears to be acceptable, come what may. Such situations are similar to moral dilemmas, and just as in moral dilemmas we will have to choose one of the (unacceptable) alternatives that come closest to being acceptable (Hansson 1999). If no available alternative is acceptable from every future viewpoint, then we should determine the lowest level of unacceptability that some alternative does not exceed in any branch, and choose one of the alternatives that does not exceed it.

The method of hypothetical retrospection is not a formula into which we can insert a set of variables, and then retrieve a definite recommendation on what to do. This may seem like a disappointment, but it is in fact what we should expect of a well-functioning decision aid for risk management decisions. The choice of risk policies is a value-based and ethical undertaking that can be affected by a wide range of human concerns. Therefore, it cannot be captured in a formula or recipe that is limited to a predetermined set of quantifiable variables.

## VII. CONCLUSION

In conclusion, we have seen that the treacherous triangle risk—science—policy can be made much more manageable than it is today. Apparent scientific controversies on risk are in many cases fictions that have been fabricated to obscure scientific consensus. Many of the genuine uncertainties can be shown to be decision-irrelevant, i.e. their resolution would have no impact on the decision at hand. And finally, ethical and decision-theoretic tools are available for dealing with the remaining decision-relevant uncertainties.

This analysis shows some of the actions that need to be taken in order to improve the science-policy interface in risk issues:

- i. We need to expose the pseudoscience that flourishes in risk issues, and put an end to its influence in policy decisions.<sup>14</sup>
- ii. We need to organize decision-making in a way that strengthens the role of independent scientific experts and limits the use of partisan expertise.
- iii. It is also necessary to systematically distinguish between the uncertainties that are relevant for the decision at hand, and those that are not. The two symmetry tests can be used to discover decision-irrelevance in cases when it is not obvious.
- iv. Ethical decision tools should be employed on a regular basis to deal with the remaining decision-relevant uncertainties.

## WORKS CITED

- Banning, Marlia Elisabeth. 2009. "When Poststructural Theory and Contemporary Politics Collide: The Vexed Case of Global Warming." *Communication and Critical/Cultural Studies* 6: 285-304.
- Barnes, Deborah E. and Lisa A. Bero. 1998. "Why Review Articles on the Health Effects of Passive Smoking Reach Different Conclusions." *Jama* 279/19: 1566-1570.
- Betsch, Cornelia, and Katharina Sachse. 2013. "Debunking Vaccination Myths: Strong Risk Negations Can Increase Perceived Vaccination Risks." *Health Psychology* 32/2: 146-155.
- Biddle, Justin B. and Anna Leuschner. 2015. "Climate Skepticism and the Manufacture of Doubt: Can Dissent in Science be Epistemically Detrimental?" *European Journal for Philosophy of Science* 5: 261-278.
- Bloch, Alan B., Walter A. Orenstein, Harrison C. Stetler, Steven G. Wassilak, Robert W. Amler, Kenneth J. Bart, Cecil D. Kirby and Alan R. Hinman. 1985. "Health Impact of Measles Vaccination in the United States." *Pediatrics* 76/4: 524-532.
- Boykoff, Maxwell T. 2008. "Lost in Translation? United States Television News Coverage of Anthropogenic Climate Change, 1995-2004." *Climatic Change* 86: 1-11.
- Boykoff, Maxwell T. and Jules M. Boykoff. 2004. "Balance as Bias: Global Warming and the US Prestige Press." *Global Environmental Change* 14/2: 125-136.
- Brandt, Allan. M. 2007. *The Cigarette Century: The Rise, Fall, and Deadly Persistence of the Product that Defined America*. New York, NY: Basic.
- Brandt, Allan. M. 2012. "Inventing Conflicts of Interest: A History of Tobacco Industry Tactics." *American Journal of Public Health* 102/1: 63-71.
- Clausen Mork, Jonas, and Sven Ove Hansson 2007. "Eurocodes and REACH: Differences and Similarities." *Risk Management* 9: 19-35.
- Collins, Harry, Robert Evans and Martin Weinel. 2017. "STS as Science or Politics?" *Social Studies of Science* 47: 580-586.
- Deer, Brian. 2011a. "How the Case Against the MMR Vaccine was Fixed." *British Medical Journal* 342/7788: 77-82.
- Deer, Brian. 2011b. "How the Vaccine Crisis was Meant to Make Money." *British Medical Journal* 342/7789: 136-142.
- Di Pasquale, Alberta, Paolo Bonanni, Nathalie Garçon, Lawrence R. Stanberry, Mostafa El-Hodhod and Fernanda Tavares Da Silva. 2016. "Vaccine Safety Evaluation: Practical Aspects in Assessing Benefits and Risks." *Vaccine* 34/52: 6672-6680.
- Dixon, Graham N. and Christopher E. Clarke. 2013. "Heightening Uncertainty Around Certain Science: Media Coverage, False Balance, and the Autis-Vaccine Controversy." *Science Communication* 35/3: 358-382.
- Doja, Asif and Wendy Roberts. 2006. "Immunizations and Autism: A Review of the Literature." *Canadian Journal of Neurological Sciences* 33/4: 341-346.
- Doppelt, G. 2007. "The Value Ladenness of Scientific Knowledge." In *Value-Free Science: Ideals and Illusions?* Edited by Harold Kincaid, John Dupré and Alison Wylie, 188-217. New York: Oxford University Press.

- Douglas, Heather E. 2009. *Science, Policy and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.
- Dreher, Rod. 2007. “Liberals’ Embrace of Science Has Been Wrong Before.” *Sun Journal*. Lewiston, ME, August 19.
- Farmer, G. Thomas and John Cook. 2013. “Understanding Climate Change Denial.” In *Climate Change Science: A Modern Synthesis: Volume 1 – The Physical Climate*. Edited by G. Thomas Farmer and John Cook, 445-466. Dordrecht: Springer.
- Francis, John A., Amy K. Shea, and Jonathan M. Samet. 2006. “Challenging the Epidemiologic Evidence on Passive Smoking: Tactics of Tobacco Industry Expert Witnesses.” *Tobacco Control* 15/4: iv68-iv76.
- Fuller, Steve. 2017. “Is STS all Talk and no Walk?”, *EASST Review* 36/1. <https://casst.net/article/is-sts-all-talk-and-no-walk/>.
- Gastaldo, Robert. A., Ivana Stevanović-Walls and William N. Ware. 2004. “Erect Forests are Evidence for Coseismic Base-Level Changes in Pennsylvanian Cyclothem of the Black Warrior Basin, U.S.A.” In *Sequence Stratigraphy, Paleoclimate, and Tectonics of Coal-Bearing Strata*. Edited by Jack C. Pashin and Robert. A. Gastaldo, 219-238. Tulsa, OK: American Association Of Petroleum Geologists.
- Guil-Guerrero, José Luis 2014. “The Safety of Edible Wild Plants: Fuller Discussion May be Needed.” *Journal of Food Composition and Analysis* 35/1: 18-20.
- Hans, Valerie P. 2008. “Jury Systems Around the World,” *Annual Review of Law and Social Science* 4: 275-297.
- Hansson, Sven Ove. 1995. “The Detection Level.” *Regulatory Toxicology and Pharmacology* 22: 103-109.
- Hansson, Sven Ove. 1999. “But What Should I Do?” *Philosophia* 27: 433-440.
- Hansson, Sven Ove. 2004. “Great Uncertainty About Small Things.” *Techné* 8: 26-35.
- Hansson, Sven Ove. 2007. “Hypothetical Retrospection.” *Ethical Theory and Moral Practice* 10: 145-157.
- Hansson, Sven Ove. 2013a. *The Ethics of Risk. Ethical Analysis in an Uncertain World*. New York: Palgrave Macmillan.
- Hansson, Sven Ove. 2013b. “Defining Pseudoscience – and Science.” In *The Philosophy of Pseudoscience*. Edited by Massimo Pigliucci and Maarten Boudry, 61-77. Chicago: Chicago University Press.
- Hansson, Sven Ove. 2014. “Agricultural Biotechnology for Health and the Environment.” In *Biotechnology and Biodiversity*. Edited by M. Raj Ahuja and K.G. Ramawat, 67-76. Dordrecht: Springer.
- Hansson, Sven Ove. 2016. “Risk: Knowledge, Ignorance, and Values Combined.” In *Moral and Social Perspectives on Ignorance*. Edited by Rik Peels, 186-204. London: Routledge.
- Hansson, Sven Ove. 2017a. “Science Denial as a Form of Pseudoscience.” *Studies in History and Philosophy of Science* 63: 39-47.
- Hansson, Sven Ove. 2017b. “How Values Can Influence Science Without Threatening Its Integrity.” In *Logic, Methodology and Philosophy of Science*. Hannes Leitgeb, Ilkka Niiniluoto, Päivi Seppälä and Elliott Sober, 207-221. London: College Publications.

- Hansson, Sven Ove. 2018a. "How Connected Are the Major Forms of Irrationality? An Analysis of Pseudoscience, Science Denial, Fact Resistance and Alternative Facts." *Métode Science Study Journal* 8: 125-131.
- Hansson, Sven Ove. 2018b. "How to Reconcile the Multiculturalist and Universalist Approaches to Science Education." *Cultural Studies in Education* 13:517-523.
- Hansson, Sven Ove. (in press). "Dealing with Climate Science Denialism: Experiences from Confrontations with Other Forms of Pseudoscience." *Climate Policy*.
- Hansson, Sven Ove and Gertrude Hirsch Hadorn, eds. 2016. *The Argumentative Turn in Policy Analysis: Reasoning about Uncertainty*. Dordrecht: Springer.
- Hansson, Sven Ove and Gertrude Hirsch Hadorn. (in press). "Argument-Based Decision Support for Risk analysis." *Journal of Risk Research*.
- Harrison, Joel A. 2013. "Wrong About Vaccine Safety: A Review of Andrew Wakefield's 'Callous Disregard'." *Open Vaccine Journal* 6: 9-25.
- Hain, Walter. 1981. *Irrwege der Geschichte. Revision der Götter-Astronauten und Atlantis*. Vienna: Hain.
- Harding, Sandra. 2008. *Sciences from Below: Feminisms, Postcolonialities, and Modernities*. Durham, NC: Duke University Press.
- Hobson-West, Pru. 2007. "'Trusting Blindly Can Be the Biggest Risk of All': Organised Resistance to Childhood Vaccination in the UK." *Sociology of Health & Illness* 29/2: 198-215.
- Hornby, D. Brock. 2014. *Judge D. Brock Hornby's 2014 Revisions to Pattern Criminal Jury Instructions for the District Courts of the First Circuit*. <http://federalevidence.com/pdf/JuryInst/1st.Crim.6.2014.pdf>.
- Kata, Anna. 2010. "A Postmodern Pandora's Box: Anti-Vaccination Misinformation on the Internet." *Vaccine* 28/7: 1709-1716.
- Kata, Anna. 2012. "Anti-Vaccine Activists, Web 2.0, and the Postmodern Paradigm: An Overview of Tactics and Tropes Used Online by the Anti-Vaccination Movement." *Vaccine* 30: 3778-3789.
- Knox, Claire Connolly. 2016. "The Human Race as Geological Agents: Communicating Climate Change Science." *Public Administration Review* 76: 974-977.
- Kramer, Bruce. 1988. "In Re Newman: The Federal Circuit Dismantles An Obstacle For Perpetual Motion Patent Applicants." *Akron Law Review* 21/1, Article 8.
- Kuske, Terrence T. 1983. "Quackery and Fad Diets". In *Nutrition in the Middle and Later Years*. Edited by Elaine B. Feldman, 291-303. Littleton, MA: John Wright.
- Larson, Heidi J., Caitlin Jarrett, Elisabeth Eckersberger, David M.D. Smith and Pauline Paterson. 2014. "Understanding Vaccine Hesitancy Around Vaccines and Vaccination from a Global Perspective: A Systematic Review of Published Literature, 2007-2012." *Vaccine* 32: 2150-2159.
- Latour, Bruno. 2004. "Why Has Critique Run Out of Steam? From Matters of Fact to Matters of Concern." *Critical Inquiry* 30: 225-248.
- Lessofo, Maurice H. 1992. *Food Intolerance*. London: Chapman & Hall.

- Lewandowsky, Stephan, Naomi Oreskes, James S. Risbey, Ben R. Newell and Michael Smithson. 2015. "Seepage: Climate Change Denial and Its Effect on the Scientific Community." *Global Environmental Change* 33: 1-13.
- Liebenberg, Louis. 1990. *The Art of Tracking: The Origin of Science*. Cape Town: David Philip.
- Liebenberg, Louis. 2013. *The Origin of Science: The Evolutionary Roots of Scientific Reasoning and its Implications for Citizen Science*. Cape Town: CyberTracker.
- Lynch, Michael. 2016. "Social Constructivism in Science and Technology Studies." *Human Studies* 39: 101-112.
- Maglione, Margaret A., Lopamudra Das, Laura Raaen, Alexandria Smith, Ramya Chari, Sydne Newberry, Roberta Shanman, Tanja Perry, Matthew Bidwell Goetz and Courtney Gidengil. 2014. "Safety of Vaccines Used for Routine Immunization of US Children: A Systematic Review." *Pediatrics* 134: 325-337.
- Matson, Dave E. 20. "How Good are Those Young-Earth Arguments?" <http://www.talkorigins.org/faqs/hovind/howgood.html> [accessed December 28, 2017].
- McKewon, Elaine. 2012. "Talking Points Ammo: The Use of Neoliberal Think Tank Fantasy Themes to Delegitimise Scientific Knowledge of Climate Change in Australian Newspapers." *Journalism Studies* 13/2: 277-297.
- Michaels, David. 2008. *Doubt is Their Product: How Industry's Assault on Science Threatens Your Health*. Oxford: Oxford University Press.
- Oldfather, Chad M. 2008. "Universal De Novo Review." *George Washington Law Review* 77: 308-365.
- Oreskes, Naomi and Erik M. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on issues from Tobacco Smoke to Global Warming*. New York, NY: Bloomsbury.
- Rudén, Christina and Sven Ove Hansson. 2008. "Evidence Based Toxicology: 'Sound Science' in New Disguise." *International Journal of Occupational and Environmental Health* 14: 299-306.
- Schafersman, Steven (1983). "Fossils, Stratigraphy, and Evolution: Consideration of a Creationist Argument." In *Scientists Confront Creationism*. Edited by Laurie R. Godfrey, 219-244. New York: Norton.
- Stiebing, William H. 1984. *Ancient Astronauts, Cosmic Collisions and Other Popular Theories About Man's Past*. New York: Prometheus.
- Trumbull, Deborah J., Rick Bonney, Derek Bascom and Anna Cabral. 2000. "Thinking Scientifically During Participation in a Citizen-Science Project." *Science Education* 84: 265-275.
- Tucker, William C. 2012. "Deceitful Tongues: Is Climate Change Denial a Crime?" *Ecology Law Quarterly* 39: 831-894.
- Wagner, Wendy E. 1995. "The Science Charade in Toxic Risk Regulation." *Columbia Law Review* 95/7: 1613-1723.
- Wazeck, Milena. 2009. *Einsteins Gegner. Die öffentliche Kontroverse um die Relativitätstheorie in den 1920er Jahren*. Frankfurt: Campus.

Weart, Spencer. 2011. “Global Warming: How Skepticism Became Denial.” *Bulletin of the Atomic Scientists* 67/1: 41-50.

## NOTES

1. The citizen science movement has largely arisen as an attempt to bridge the perceived gap between science and everyday knowledge. (Trumbull *et al.* 2000)

2. American juries are often instructed that they have “a duty to find the facts” and “must not be influenced by any personal likes or dislikes, prejudices or sympathy” (Hornby 2014, 60). Juries have similar functions in many other countries (Hans 2008).

3. For a few examples of this argument, see Kramer (1988, 129), Dreher (2007), and Box (2016, 975). For an analysis of the argument, see <http://www.critical-thinking.org.uk/critical-thinking/bad-arguments/science-has-been-wrong-before.php>.

4. If the conflict is not (genuinely) scientific, then the uncertainty is not (genuinely) scientific.

5. For other approaches to such mechanisms, see for instance Lewandowsky *et al.* (2015) and Farmer and Cook (2013). For documentation of the claims of science denialists, see for instance Kata (2012), Farmer and Cook (2013) and Hansson (2017a).

6. ‘Science in a broad sense’ differs from the standard English usage of ‘science’ in also including the humanities. It corresponds to the usage of the Latin *scientia* and the German *Wissenschaft*. See Hansson (2013b, 62-66) for a justification of the use of the broad rather than the narrow concept of science in this context.

7. The distinction between fabricated controversies and legitimate scientific dissent is essentially one of scientific quality and willingness to take arguments. On this distinction, see Hansson (2017a), Hansson (in press), and Biddle and Leuschner (2015).

8. It can however be argued that by understating the uncertainty they make pseudoscientific statements about the state of knowledge about the risk.

9. The description of this case is based on Hansson (2016). The Wakefield hoax is used here only as an example. For more general discussions of anti-vaccine activism and vaccine hesitancy, see Kata (2012) and Larson *et al.* (2014).

10. This is usually described as a comparison of the risks and benefits of vaccination. The reformulation in terms of two classes of risk is preferable here since the symmetries of the situation comes out more clearly.

11. Perhaps more importantly, one of the complications of measles is subacute sclerosing panencephalitis, a deadly disease with symptoms including personality changes and progressive mental deterioration that ends up in a vegetative state.

12. The test can also be applied to uncertainties concerning potential benefits. Then we have to consider whether it can be outweighed by some potential risk that has equal or larger moral weight.

13. It was first presented in Hansson (2007) and further discussed in Hansson (2013a, 61-73).

14. See Hansson (in press) for a discussion on how this can be done.

# *Risk Impositions, Genuine Losses, and Reparability as a Moral Constraint*

Madeleine Hayenhjelm  
Umeå University, Sweden

**ABSTRACT.** What kind of moral principle could be sufficiently restrictive to avoid the kind of large-scale risks that have resulted in catastrophe in the past, while at the same time not be so restrictive as to halt desirable progress? Is there such a principle that is not merely a precautionary principle, but one that could be based on firm moral grounds? In this article, I set out to explore a simple idea: might it be the case that *reparability* could serve as a moral constraint against risky policy decisions? The idea is simple, but comes in two forms. First, that it is morally wrong to impose a risk for harm that is, in principle, irreparable, such that it would bring about a permanent loss of a kind *qua* kind. Second, that it is morally wrong to impose a risk for a reparable (but not compensable) harm that exceeds what realistically could be repaired. I set out here to do two things. First, I describe the moral problem to be addressed for any principle-guiding decisions about risk. A central claim in this article is that risk decisions are epistemically impaired decisions, and that we must, alongside outcome uncertainty, also take both epistemic uncertainty and moral uncertainty into account. Second, I introduce the idea of reparability as a moral constraint in the form of two versions of the Reparability Principle. Such a principle, I argue, could have some interesting advantages that seem both morally intuitive and that come with some advantages against some of the epistemic challenges posed by risk impositions.

**KEYWORD.** Risk, uncertainty, genuine loss, reparation, reparability

## I. INTRODUCTION

**I**n the present contribution, I set out to explore a simple idea: might it be the case that *reparability* could serve as a moral constraint against risky policy decisions? The idea is simple, but comes in two forms. First, that it is morally wrong to impose a risk for harm that is, in principle, irreparable, such that it would bring about a permanent loss of a kind *qua* kind. Second, that it is morally wrong to impose a risk for a reparable

(but not compensable) harm that exceeds what realistically could be repaired.

Russian roulette examples have become standard illustrations of the moral problem of risk.<sup>1</sup> However, actual risk decisions differ in two important ways from such idealised cases: they differ in terms of their epistemic outset conditions, and they differ in the nature of their impact. In the Russian roulette case, as in all cases of ‘decisions under risk’, we know all the potential outcomes and we know the probabilities for those outcomes. Furthermore, risk in the idealised Russian roulette examples are ‘pure risk impositions’, as per Thomson’s distinction (1986), where, if the anticipated harm never occurs, no harm has been done.<sup>2</sup> The more pressing real cases of risk differ from the Russian roulette cases in both of these respects. In contrast to the idealised cases of Russian roulette, most cases of actual risks are not cases of ‘pure risk impositions’ but are cases of ‘mixed risk impositions’, such that they both cause a harm and impose a risk for further harm (Thomson 1986). According to Thompson:

Suppose A shoots B in the stomach. Then A causes B a very serious harm; but he also imposes a risk of still more serious harm, namely death [...] B has a twofold complaint against A: he has grounds for complaint against A that A caused him serious harm, and he has yet another complaint against A in that A has imposed a high risk of death on him (1986, 175).

A number of considerations make moral decisions about risk different from other moral decisions. First, there is *harm* that may or may not come about. This makes the considered action more morally problematic than a straightforwardly safe option, but less morally problematic than an action that is straightforwardly harmful. Any moral principle for risk decisions must be sensitive to this balance and guarantee that trivial risks are not made impermissible or serious risks made permissible. Second, we are in an *epistemically weaker position* than we would have been had we pursued an action known with certainty to cause harm or an action

known with certainty to be safe. This goes beyond not knowing which one of a set number of outcomes will occur. In some cases, it means that we will simply know too little to be able to make a well-informed decision.

I take it that five parameters determine the urgency and seriousness of any risk: (i) the degree of its impact being human-made or human-preventable; (ii) the magnitude of harm (its reach in time, global impact, and size of population); (iii) the degree of irreversibility of harm; (iv) the probability of it occurring; and (v) the degree of uncertainties involved. The first is merely a prerequisite for the urgency and moral dimension of the risk. The risks we are morally concerned with are the risks we can do something about. The second and third aspects relate to the potential outcome and the degree of harm that can be caused. The fourth aspect refers to the probability for such outcomes to occur. And the fifth aspect refers to the degree of uncertainty and ignorance about probability, risks, and outcomes. Presumably, the risks we are most likely to regret in the future are those that could cause harm of great magnitude in impact, scope, and/or irreversible damages. The risks most difficult to prevent are those that we impose in ignorance or under great epistemic uncertainty, such as new risks, risks with far-future impact, or risks with systemic impact. The challenge arises when the degree of ignorance is not known and when such ignorance may ‘hide’ risks for catastrophic outcomes.

What kind of principle could possibly guide such risk decisions? My aim here is to search for a principle sufficiently restrictive to avoid catastrophic outcomes but not so restrictive as to halt desirable progress.

## II. THE EPISTEMIC PROBLEM OF RISK IMPOSITIONS

The literature on risk decisions, whether in moral philosophy, applied ethics, or decision theory, tends to focus on two points in time: the time of decision and the time of consequence.<sup>3</sup> Let us refer to these times as T1 and T3, respectively. The challenge is to make the right decision about

what to do at T1 when given a particular aim, the success of which we will know only at T3.

Typically, the epistemic property of the decision at T1 is either framed as a *decision under risk* or as a *decision under uncertainty*. If it is a case of decision under risk, then the moral agent would know the relevant options, the relevant possible outcomes, and the probability for each outcome for each option. If it is a case of decision under uncertainty, then the moral agent would know the relevant options and the relevant possible outcomes, but not the probability for each outcome. Most of the philosophical literature on risk tends to use ‘risk’ broadly to cover both cases. But, as Hansson (2009) has pointed out, in most real-life cases we tend to know much less than in either of these two cases.

From an actualist consequentialist approach, the morally significant time is T3; only at T3 can we know whether a given decision was right or wrong. Other more prospective versions of consequentialism would still take the consequences at T3 as central, but would shift the focus to how these are rationally perceived at T1.<sup>4</sup> From a traditional deontological point of view, T1 is the morally most significant point in time. Moral principles based on T3 have the advantage that moral rightness or wrongness will perfectly match how things went. Thus, a catastrophe will always – or in most cases – also be morally wrong. The drawback is that such principles may not be sufficiently action-guiding to avoid such actions. Moral principles based on T1 may have the advantage of being action-guiding, but they are insensitive to how things play out after the decision has been made. Either way, the standard view, with the exception of virtue ethics, tends to be that moral rightness is determined by either T1 or T3. We will either know at T1 if the decision was morally justified (given the reasons, principles, etc.), or we will find out if it was morally right at T3.

In this section, I wish to draw attention to the time between these two – namely, T2 – and the fact that things can change in this intermediate time in morally significant ways. If so, T2 may also be a time for

possible re-consideration and additional action. When it comes to actions with immediate effects, as in the Russian roulette cases, the significance of T2 is minimal. The time between the decision and the outcome would be a matter of seconds. But when it comes to new risks and risks with far-future effects, the moral significance of T2 is central. What we know at T1 and what we could learn over the long stretch of time comprising T2 and up until the final consequence at T3 could change dramatically. In particular, the most catastrophic outcomes seem often to be the result of risks not foreseen at T1 and only realised later. Consider the case of arsenic poisoning in Bangladesh (see, for example, Smith, Lingas and Rahman 2000). Millions of tube-wells were built in villages around the country in order to provide clean drinking water to combat water-carried diseases and child mortality. Two decades later, a massive problem of arsenic poisoning was discovered. The concern here is that none of the standard principles that guide moral action would have advised against the initial line of action.

At T1 there are a number of things that the moral agent may not know. Three kinds of ignorance or uncertainty are particularly problematic: *outcome uncertainty*, *epistemic uncertainty*, and *moral uncertainty*.

In the simplest case of *outcome uncertainty*, we know all there is to know about each of the possible outcomes; the only thing we do not know is which one of the possible outcomes will, in fact, occur. In such cases, the outcomes are limited to what we could refer to as *primary effects*, or the effects intended. In most real cases of risk, we are more concerned about *secondary effects* in the form of unintended side effects. Take DDT (Dichlorodiphenyltrichloroethane) as an example. The problems with DDT did not arise from its anticipated primary effect: it did have the consequences expected of a pest-control agent. The problems came in the form of side effects. But primary effects and side effects can, to some degree, both be anticipated. Sometimes the risks do not arise from the activity itself by itself but arise from *external risks* that trigger, alter, or amplify the effects. The Fukushima accident can serve as an example here. In that case, there

was both a tsunami and an earthquake that affected the power plant. In other cases, the outcomes we worry about are the *contributory effects* to larger systems. This is the concern that has been raised about combining gene-editing of insects and gene drives and its potential effects on entire ecosystems (see, for example, Lunshof 2015):

Once introduced, these genetic changes are self-propagating. If released beyond the laboratory, the effects would spread with every new generation and would quickly run out of control (Lunshof 2015, 127).

Lunshof further argues that “[...] the broader ecosystem disruption is unknown and would require extensive mathematical modelling to estimate” (2015, 127).

The second type of uncertainty is *epistemic uncertainty*. At T1, we may not know how much of the relevant information we do or do not know. It may be the case that, at T1, we actually do not know enough to make a particular decision, but that we only think we do. Not knowing whether or not this is the case is to act under some kind of *epistemic risk* (Sahlin and Persson 1994).<sup>5</sup> Consider the epistemic situation of a hypothetical Ministry of Transportation at the dawn of motoring in the 1910s-20s. Mariscal and Petropanagos (2016) entertain just such a thought experiment in a recent paper comparing CRISPR with the impact of the Ford Model T. They ask what risks could such an agency have foreseen, what risks they would have missed, and what risks they would have been concerned with that we no longer see as that important. They suggest that demand for oil as a driving geopolitical factor, climate change, and air pollution are risks that they would have been likely to miss. Instead, they would probably have foreseen negative impacts on the horse industry and on railroad traffic, traffic planning, and the like.

The third type of uncertainty is that of *moral uncertainty* (see Lockhart 2000). This has been discussed in some of the philosophical literature but not very much in applied contexts. The basic idea is that we may not

know which moral theory is the right one. This becomes particularly problematic when two theories are equally subjectively probable, but are mutually incompatible, giving conflicting guidance on what is the right thing to do (see, for example, Sepielli 2013).<sup>6</sup> Here I use the idea in a much broader sense: we may not know for certain at T1 that our current moral considerations are the right ones, and the values upon which we base our decision may seem clearly morally wrong by the later generation that would be affected by that decision. As an example, consider the idea that we would come to regard all mammals as our moral equals in the future. Were this to be the case, then we would most likely come to re-evaluate a number of decisions in medical research that we, perhaps, currently believe to be morally right. The general point is that many of the risk decisions that we make today could be such that they fundamentally change the background conditions for future generations. If the moral values such a world is built upon are perceived as repulsive to future generations, this could lead to a catastrophe, albeit of a moral kind.

Apart from what is known, and not known, at T1, the epistemic situation could change in a number of ways across T2, but prior to T3. Some ways in which the epistemic situation could change might warrant a complete re-assessment of any decision made at T1. Consider examples like the following:

- i New alternative options may arise.
- ii New risks arising from the action pursued may be learned.
- iii New details about outcomes may arise (such as new side-effects, new knock-on effects, etc.).
- iv New sources of external risk factors that could impact the risks and outcomes of the action pursued.
- v New populations not previously thought to be at risk might be affected.
- vi Moral standards and values may have changed.

Any one of the epistemic changes listed above could make a moral decision made at T1 seem like the wrong decision. In other words, what we do not know or wrongly believe that we know at T1 could change across T2 such that new information changes the moral outlook.

### III. THE MORAL PROBLEM

Let us now restate the initial questions against the background of the epistemic problem discussed in the previous section. What kind of moral principle could possibly be sufficiently restrictive to avoid the kinds of risk that have resulted in catastrophe in the past while at the same time not being too restrictive as to halt all progress? Is there such a principle that is not merely a precautionary principle, but one that could be based on firm moral grounds?

It has turned out to be surprisingly difficult to get straightforwardly satisfactory answers from normative theories about the moral justification of risk impositions (see Hayenhjelm and Wolff 2011). The problem derives from the fact that imposing a risk of harm may be morally less bad than imposing harm for certain, on the one hand, and the fact that imposing a risk of harm may be morally worse than not imposing any risk at all, on the other. This makes any straightforward application of standard moral principles developed for actual and certain harm, rather than for risks, problematic. This basic problem has been addressed in various ways. The greatest challenges come with the various uncertainties, including epistemic and moral uncertainty, discussed in the previous section.

#### *Two Favoured Explanations*

Two frameworks have come to dominate some of the ethics of risk literature. These are consequentialism (particularly in a cost-benefit version) and rights-based frameworks.<sup>7</sup>

Consequentialism initially seems to get the moral problem right (see, for example, Thomson 2001, 6-9). The reason we are concerned about risks is that we are concerned about outcomes and states of the world. It seems exactly right that what makes risks morally problematic has to do with its consequences. Furthermore, at the very heart of consequentialism lies the balancing of benefits against negative outcomes: the claim that risks cannot be avoided altogether, but must be balanced against prospective goods comes naturally to consequentialist responses. Consequentialism also allows for proportionate discounting of harm according to the probability for harm; it can thus avoid treating merely possible harms as actual harms or not as harms at all. For these reasons it is not hard to see that cost-benefit analysis, as a form of applied economic consequentialism, aiming to maximise benefits over costs, has become a dominant framework in risk policy literature. It also aligns well with the rational framework in decision theory, which, in much the same way, is also based on maximising outcomes.

The epistemic element is much more difficult to solve, however. The rightness or wrongness of an action depends ultimately upon its consequences, and these may not be known. This problem increases the greater the epistemic uncertainty.<sup>8</sup> The problem is roughly this. Either the actual consequences or the expected consequences determine right and wrong. If it is actual consequences, then consequentialism cannot be very action guiding when it comes to decisions taken under great uncertainty or great epistemic risk. In the case of expected consequences, consequentialism would be action guiding, but would seem to lose some of its moral appeal in connecting right and wrong to actually making the world better or worse. My worry is this: many of the risks imposed that in hindsight turned out to be catastrophic may very well have been a result of the most ‘rational’ choice according to consequentialist calculations based on what was known at the time. Furthermore, by giving priority to maximisation, there is the risk of undervaluing precautionary measures that cannot be epistemically justified at the time. In other words, relying on

numbers based on the risks that we foresee does not guarantee that we prevent catastrophes or great harm.

Most of the worries raised against consequentialism have to do with aggregation. It has been argued that consequentialism cannot fully do justice to issues of fairness or individual rights, since all harms and benefits are aggregated across individuals. We could also add to this the problem of *discounting* future harm in relation to cost-benefit analysis. If all risks are treated as future monetary losses, then these will be discounted much in the same way as all costs and gains in the future are discounted in the same model. This means that the further one moves into the future the less any possible harm counts (see, for example, Weisbach and Sunstein 2009; Broome 2008; Dasgupta 2008).

A further consequence of morality viewed as a net balance is that all negative outcomes are, in principle, *compensable*.<sup>9</sup> This is the most obvious in cost benefit frameworks where all outcomes are, as far as possible, monetised. There is a price for every risk: a level of compensation that would outbalance the harm of the risk by the good of the benefit.<sup>10</sup> If a person P suffers harm from a certain risk, then we could either remove the harm or offer something else of equivalent value, and that would be, morally speaking, equivalent. In standard monistic versions of consequentialism, the same logic holds: a loss in one part of the system can be outweighed by a gain elsewhere in the system.

Usually, such worries are met by appeals to pluralism and the incommensurability of values. But this need not solve the fundamental problem. There is nothing *in principle* uncompensable, on this view, such that a greater benefit of some commensurable kind could not compensate for it. This lack of any absolute barrier does not seem to fully capture the severity of catastrophic risks.

All three of these worries – about aggregation, discounting, and compensability – stem from one and the same feature of the theory, which we might refer to as the abstraction of value from the states of affairs.<sup>11</sup> This abstraction is an essential part of the theory since maximisation

would not be possible without being able to compare benefits and costs with each other and across states of affairs. If we can abstract the value away from a particular state of affairs and treat that value as separate from its material basis, then all things would be compensable in some sense, at least for as long as similar values can be found to replace it.

The basic idea of rights-based ethics when applied to risk typically comes in the form of the right of an individual to not have risks imposed upon them (McCarthy 1997). Thus, it is always morally wrong to impose a risk of harm on another person unless he or she has given his or her consent to such a risk. A moral theory based on individual rights makes it very clear precisely what makes abuse, exploitation, and harming of others wrong. Instead of abstracting and aggregating values across states of the world, it draws a moral line of impermissibility based on the rights of individuals. It thus has an advantage over consequentialism insofar as it can make sense of the idea that certain types of actions are simply not permissible because of the kinds of actions they are. The main drawback is that this barrier is unnecessarily anthropocentric and that it does not seem able to capture the scope and magnitude of catastrophic harm.

Rights-based ethics have strong anthropocentric leanings: risks are only morally problematic to the extent that they affect the rights of individuals. These could be expanded to include animals or possibly even ecosystems, but this could give rise to serious conflicts of interests. The fundamental worry remains: without claimants there is no rights violation and hence no wrong. Should there, for some reason, not be any individuals left to have rights infringed at the time of consequence, then no one's rights would have been violated. The implication is that, even if we destroyed the whole planet, this would not be impermissible should no rights be violated as a consequence of this act. We could perhaps imagine that all of humanity had left earth before this occurred, or that humanity was wiped out by some deadly disease or by a meteorite, or that human nature in some distant future was so genetically different that none of our rights applied. Similarly, should all affected individuals somehow give

their consent to such destruction, then it would not be impermissible. We could, perhaps, imagine that everyone was convinced of the romantic value of being the last generation, or that going extinct somehow was the perfect moral atonement, and thereby, on one or more of these bases, give their consent.

When it comes to potentially catastrophic risks with global impact, it seems that rights-based ethics cannot explain the magnitude of the risk or the moral concern. What makes catastrophic risks wrong, according to the theory, is that they violate the rights of individuals. Such risks do not constitute a graver moral wrong than any other kind of rights violation. It will presumably affect many individual rights (of currently living and future individuals) but it is not clear that the theory would suggest that we add these together. Any violation of rights is a serious moral wrong and thus impermissible, but there is not really anything else left in the tool-box capable of capturing any other kind of wrong that is *more* serious than that. The worst outcome is the same as the lesser outcome: it violates a right. Even if we could add all individual complaints into a collective complaint, this does not solve the worry about anthropocentrism.

#### IV. A NEW ATTEMPT: THE REPARABILITY PRINCIPLE

Here is a bold suggestion: a risk imposition is morally wrong if the possible negative outcome cannot be reversed, repaired, or replaced. The worry that risks that cannot be reversed, repaired or replaced raises is not merely a worry about magnitude and probability. It is also about causing harm that cannot be undone, healed, or repaired in time to prevent a catastrophe. Contrast such irreparability with full restoration of natural states of affairs or objects. Something broken can be repaired by having its broken parts re-assembled or fixed. Something broken or lost can be replaced if there is something of the same kind that could replace all the values lost. A state of affairs is reversible if it can be returned to some previous stage before damages and if that previous stage can be regarded

as in some sense valuable and un-broken. The suggestion here is that that the imposition of a risk of harm that cannot be restored in any of these ways it is morally wrong. The idea is that there is something fundamentally wrong in permanently making the world a worse place. There seems to be something morally significant about losses or damages that are irreversible in the sense that, once something is gone, it is gone forever and will never come back. The idea would look something like this:

*The Reparability Principle (first attempt):* All else being equal, it is always morally wrong to impose a risk of harm such that that harm cannot be reversed, repaired, or replaced.

The exact phrasing could be polished in various ways, but I take the main idea to be fairly intuitive: there is a moral limit to our freedom of action set by what can be repaired. ‘Harm’ is intended in a very broad sense to cover damages, losses, and any detrimental impact on the state of the world (in a natural sense). ‘Morally wrong’ is intended in a broad sense to point to something that always comprises a compelling reason against doing something on moral grounds. Let us turn our focus, for now, to reparability and the idea that this is of moral significance.

Let us distinguish between three kinds of loss: *simple loss*, *reparable loss* and *genuine loss*. A simple loss is a loss that is compensable and replaceable. It can be replaced or traded for something else.<sup>12</sup> A tulip that has withered can be replaced with a new one or it can be fully compensated for with money sufficient to cover the cost of a new one. Such a loss is a simple one. A genuine loss, by contrast, is a loss that remains over time and that cannot be compensated for by something else. It constitutes a permanent loss and leaves a ‘void’, as it were. Between the two are reparable (genuine) losses. These are such that they can, in principle, be repaired or replaced, but which, if they are never repaired, become genuine losses. We could thus distinguish between two different versions of the Reparability Principle, one excluding

genuine losses and the other excluding excessive losses that make reparations unfeasible or unlikely.

*The Reparability Principle (against genuine losses):* All else being equal, it is always morally wrong to impose a risk of genuine loss for what could not even in principle be repaired.

*The Reparability Principle (relative):* All else being equal, it is morally wrong to impose a risk for uncompensable but repairable losses that are greater than what can realistically be fully repaired or replaced.

Again, these principles are only meant to apply to uncompensable losses (however that is defined). Thus, if we assume that some losses are such that they cannot in principle be compensated for or replaced with something else, such as money, then we must never cause damages such that we cannot repair or replace them with something of the same kind. That is the general idea. Now let us investigate these two versions of the principle in turn.

### *Genuine Losses and the Reparability Principle*

What makes the imposition of a risk for in principle irreparable genuine losses wrong? The reason we care about risks is that we care about outcomes. We care about whether or not we make the world a worse place. If it is morally problematic to leave the world a worse place than we found it, then it would be even more problematic to make the world permanently worse.

The basic intuition behind both above principles is this: it is morally wrong to stake what cannot be repaired and to destroy what cannot be made whole. The idea is roughly one of economy or natural limits. It gives a kind of rationale for why blowing up a world heritage site, destroying the last of a species, destroying a whole culture, exterminating the whole of mankind, and the like, are particularly problematic. It is

considerably worse to kill the ten last humans than to kill ten humans. The point here is that this difference is not merely one of magnitude but, morally speaking, involves a different kind of action. However, it could be argued that irreparability is not significant by itself, but is only significant as an ‘amplifier’ of magnitude. This is roughly Sunstein’s position (2010, 235) when it comes to ‘irreversibility’ and ‘irreversible damages’. But the worry could also be raised against *genuine losses*. As Sunstein says:

Environmentalists who are concerned about irreversibility have something more particular in mind. They mean something like a large-scale alteration in environmental conditions—one that imposes permanent, or nearly permanent, changes in those conditions. Any such alteration might justify significant concern, but it should be clear that irreversibility in this sense is not a sufficient reason for a highly precautionary approach. At a minimum, the irreversible change has to be for the worse, and it must also rise to a certain level. A truly miniscule change in the global temperature, even if permanent, would not justify expensive precautions if it is benign or if it imposes little in the way of harm. For this reason, it is tempting to understand the idea of irreversibility as inseparable from that of seriousness. A loss of a wisdom tooth is irreversible, but not a reason for particular precautions on behalf of wisdom teeth; a loss of an extremely small forest, with no wildlife, hardly justifies a special principle, even if that loss cannot be reversed. A loss of a large forest, with ample wildlife, is a different matter. At first glance, then, irreversibility matters only because of its connection with the magnitude of the harm; irreversibility operates as a kind of amplifier (2010, 235).

I believe that Sunstein is right that irreversibility must be about changes for the worse. I shall prefer the term ‘irreparable’ over ‘irreversible’ for this reason.<sup>13</sup> However, I also believe that permanence is morally significant, and not only when it applies to an already serious harm. Take the most trivial of losses: a glass of milk is dropped such that the milk is spilled and the glass is broken into pieces. Now let us assume that P imposes a risk for such a harm on Q. P balances a glass of milk on her head in Qs kitchen. P drops it and pieces of glass spread on Qs kitchen floor and milk spills everywhere. This can quickly be wiped up and the

pieces of glass removed such that a few minutes later there is no trace of the harm, save from the loss of milk and the glass, both of which could easily be replaced. Now, consider the same case but with the difference that it is, in principle, impossible to ever remove the pieces of glass or wipe up the milk. In fact, they will both remain on the floor until the end of time. This would no longer be a trivial harm. Q could no longer have guests over in the kitchen or let children play there. Large parts of the floor would be unusable and it would become difficult and hazardous to move around. Over time, the milk could attract all sorts of insects and start to smell. The milk and the glass pieces might cause serious injuries due to slips and cuts. It could, of course, be argued that the permanence is still a mere amplifier of the original harm, but it seems that the seriousness of the harm stems from the permanence more than it does from the magnitude of the harm. It is the pervasiveness and inescapability of the harm that makes it morally problematic. It seems to me that the underlying intuition behind the worries about ‘irreversible damages’ in terms of the permanence of the loss, or, as in the above example, the permanent addition of harm, is highlighting something important.

Killing ten innocent humans would be wrong on any moral account. However, there is something fundamentally worse about killing the last ten humans. Killing those ten humans would cause all of humankind to perish forever. The loss of the last of a kind constitutes a particular kind of irreparable harm where the whole material basis for a future possibility of a certain kind of valuable experience can no longer be had. The core moral idea behind the Reparability Principle, in the genuine loss reading, could thus be phrased in the following way: all else being equal, it is morally wrong to contribute to the complete and permanent removal of the material basis for a certain type of manifest value. Such a loss exceeds the sum of the particular losses. They constitute a loss of a kind *qua* kind. Such losses are always genuine losses (unless compensable). Such a loss of a kind could occur in two ways. It could be the *local* loss of a kind or the *metaphysical* loss of a kind.

*Metaphysical Irreplaceability:* An object lost or damaged beyond repair is metaphysically irreplaceable if and only if no object of that same kind exists or ever will exist.

*Local Irreplaceability:* An object lost or damaged beyond repair is locally irreplaceable if and only if no object of that same kind can ever come to exist in the same place as the one lost.

A local loss of a kind would be a glacier that dries up in the Andes and results in the loss of the only fresh water available to all dependent communities. A single island that is submerged under water and so lost forever would also be a local loss of a kind. But the distinction is not strict. Is the loss of the last ten humans on earth a metaphysical loss or merely a local loss? In both cases the losses are grave enough to be in principle irreplaceable and irreparable.

Let us make the following assumption: it is, all else being equal, always better to increase the number of positive potentialities in the world. ‘Potentiality’ may seem somewhat Aristotelean and archaic, but in the context of risk it could be useful. (The idea of being able to count these potentialities may seem even more curious.) However, if adding the number of potentialities for the good is a good thing, then, according to the same principle, removing the number of potentialities would be a bad thing. We would then get something like the following:

*Principle of positive potentiality:* All things being equal, it is always morally good to increase the number of possibilities for the good in the world.

*Principle of negative potentiality:* All things being equal, it is always morally bad to remove the number of possibilities for the good in the world.

If these two principles track valid moral concerns, and, if the causing a genuine loss of a kind would violate the principle of negative potentiality, then we would have a kind of argument in support for the Reparability Principle.

*Reparable Losses and the Relative Reparability Principle*

Some damages and losses are not, strictly speaking, irreplaceable or irreparable; they are only very unlikely or even impossible to replace or repair *given the circumstances*. They may not be reparable or replaceable due to social, political, or financial circumstances. Alternatively, the relevant competence necessary for such repairs may not exist. Had those circumstances been different, or were they to change in the future, the same damages and losses would be reparable and replaceable. Let us refer to such losses as *circumstantially irreparable losses*.

*Circumstantial irreparability:* An object lost or damaged is circumstantially irreparable if and only if the circumstances make it impossible or extremely unlikely to either repair or to fully replace that loss.

We could, for instance, imagine some environmental disaster in a very poor nation. None of the damages may be irreparable in principle, only very costly and difficult to repair, and there may be neither any possible resources to pay for such reparations nor the know-how required to make them. A risk for a circumstantially irreparable loss is only conditionally impermissible as per the relative version of the Reparability Principle. They are only impermissible, all else being equal, to the extent that the damages are unlikely to be repaired (and conditional on other moral concerns). In other words, should these circumstances change, the very same risk might then be permissible.

However, merely predicting or knowing that certain damages are likely to be repaired would not by itself make them permissible. First of all, other moral considerations may need to be met. But more worryingly, predicting the possibility of repair does not seem to be what is morally relevant. That a risk-imposing person P calculates on Q an obligation to repair any damages that P brings about does not make P's risk-imposition permissible. It is tempting to instead make the permissibility of P's risk-imposition conditional on P's commitment to repair any damage that the

risks P imposes cause. But this would be too restrictive. For one, it would make all long-term risks that could have impacts beyond one lifetime impermissible. The reasonable requirement must fall somewhere between merely predicting reparability and personal responsibility for reparations. We would need something like a ‘vouching’ for the realistic possibilities of repair. The idea is that, as far as possible, a risk-imposing agent would seek to repair any damages caused or would in other ways ensure, through planning and preparations, that such reparations are realistic. Perhaps we could reword the Reparability Principle in the following way:

*The Reparability Principle (responsibility):* All else being equal, it is morally wrong to impose a risk for uncompensable but repairable damages that exceeds what the moral agent can vouch for as being realistically likely to be fully repaired or replaced.

The Reparability Principle would then need to be paired with something like the Responsibility Principle in order to exclude merely depending upon the reparations of others.

*The Responsibility Principle:* All else being equal, it is morally required, as far as is possible, to seek to repair any repairable loss caused or to replace any replaceable loss caused.

Together, these two principles – the Responsibility Principle and the Reparability Principle – suggest the more basic idea of a natural boundary for risking. The ‘size’ of permissible risks is not set independently, but is set relative to what can be repaired. The more things that can be repaired, the more things may be risked. What these two ideas bar against is gambling with risky outcomes such that we could end up with large scale damages without the necessary plans, skills, resources, and competence to repair such damages. Instead, it seeks to ensure, as far as possible, that, no matter the kind and size of the risk, should the worst happen, it can be repaired. This means that there are three moral obligations with regard to

risk: (i) to avoid, as far as possible, risks that cannot be repaired; (ii) to ensure the realistic possibility for reparation of damages; and (iii) to only impose risks that we have reason to believe can be fully repaired. The point is that the third obligation depends on the second: the more that we make sure can be repaired, the more we can risk. Risks can then be made more or less permissible as a result of deliberate choices and policies.

## V. REPARABILITY AND THE EPISTEMIC PROBLEM

Let us now return to the epistemic problem of moral risk decision and see if there are any epistemic advantages to the Reparability Principle compared to the other theories discussed. There are two aspects I will focus on: the *epistemic demandingness* of a moral principle for risk decisions and its *epistemic flexibility* such that it can adjust to the fact that the situation may change over time.

The epistemic problem of risk impositions will not go away, regardless of the moral principle employed. But what we will need to know in order to determine whether an action is morally right or not differs. In order to apply the traditional utilitarian criterion of rightness, we would need to know if an action would maximise positive outcomes over negative ones and whether it does so more than any alternative action would. This is epistemically very demanding. Applying deontological frameworks that determine the rightness of an action based on the reasons, principle, and/or procedures that led to the decision are not epistemically burdensome in the same sense. Their weakness lie in the inflexibility that follows: if a decision made at T1 was the morally right decision then this would be true no matter what follows. How demanding rights-based frameworks are depends on what the relevant rights are: that is, whether they are primarily thought to be violated by the mere act of imposing a risk or if they are thought to be violated by the consequences of that risk (or both). Additionally, determining moral rightness based on individual rights becomes very uncertain if we need to take the rights of individual future

rights-holders into account. Perhaps the greatest worry about rights-based ethics, when applied to risks for future consequences, is one about vagueness and the number of assumptions that need to be made in order to reach any conclusions about what may or may not be done. Hansson (2009) has argued that, epistemically speaking, risk decisions are more like a walk in the jungle than a calculated game of roulette at the casino. If we borrow this analogy, we could argue that a traditional consequentialist would need to know, at every point, if the intended path was comparatively the best one based on all the dangers that could ensue for each alternative path. This is epistemically daunting. The deontologist would need to know that the reasons for the chosen path complied with moral duty or some other moral criterion. Given that the moral criterion is known and the reasons are transparent, this would not be very epistemically demanding. To decide on a path from a rights-based point of view, we would need to know to what extent each path could violate the rights of any potential rights-holder we might encounter in the jungle. So what about a person making the decision based on our Reparability Principle? Such a person would refrain from imposing risks for destroying the whole jungle, or killing any animal that could be endangered, or polluting any waterways that may be unique. That is what the first principle would require. The second principle would not require any detailed predictions and comparisons about the paths, but it would require well-informed confidence that the companions selected and the tools and the equipment packed would suffice to deal with any danger encountered.

How *epistemically demanding* is the Reparability Principle? Let us identify what we need to know in order to apply the first and second versions of the principle. In order to apply the Reparability Principle against genuine losses, we would need to have some kind of reason to think that a certain risk could lead to genuine losses. The epistemic challenge would perhaps be centred not so much around whether we know something is the last of a kind, but whether we know that certain future actions might causally lead to such a loss. It provides us with a moral principle to apply

if we cannot exclude such risks, but it would require us to know that there could be such a risk. In order to apply the second version of the principle, we need to know if the possible outcomes fall within what can realistically be repaired. It is this version of the principle that I think has the strongest epistemic advantages. We need not know exactly what the outcomes will be, how they will occur, or to whom; we only need to know how well prepared we are to deal with such outcomes. It is a much lighter epistemic requirement to know what is within one's (collective) capacities to do than it is to predict all the ways in which events may go wrong and whom they might affect.

In other words, we may not need to know all the details of all possible outcomes to know whether or not we could repair the harm if the risks were familiar and low-scale. A skilled artisan may not have predicted all the mistakes he could make, but he would have all the tools, education, and experience to be confident that any such mistake could be fixed. In most cases, such a person would be right. In the same way, a government could make sure that there is the relevant expertise, the relevant research, the relevant competence, and the relevant resources set aside to deal with risks falling within the normal range. For novel risks this would not suffice. Decisions regarding such risks would need to proceed in steps, such that each step was known to be reparable and that no step exceeded what was, at the time, known to be reparable.

Let us turn to *epistemic flexibility*. Previously, I argued that a central problem for risk decisions is that they are made under an epistemic impairment of one kind or another. Apart from outcome uncertainty with regards to direct or secondary effects, there could be unknown risks that lead to a potential catastrophe. Furthermore, there is the problem of moral uncertainty: the values we based our decision on may change. In addition, we may learn along the way about much better alternatives that were not available at the time of the decision. Two aspects seem particularly relevant to determine a moral principle's epistemic flexibility in the relevant sense: (i) the degree to which an unexpectedly dangerous course

of action can be refrained from with minimal harm, and (ii) the degree to which moral rightness is not determined at the time of decision.

Consequentialism, at least of the actualist variety, would not fix the rightness of an action until after the event. This means that, should evidence emerge that the current path could lead to great catastrophe, the decision ought to be changed. However, the aim of maximising outcomes could lead to greater risks taken based on what was known, making such a change potentially very difficult or costly. If we have no known reason to spend money on a precautionary measure at T1, and if the same money could be invested in something expected to bring about great benefit, then we ought to spend it in the most beneficial way. In other words, the degree of harm to repair something if it is discovered too late could be extensive. Furthermore, any re-calculation of the best alternative action at T2 might be as epistemically demanding as the first decision, given that we must still act in a way that is expected to maximise the outcome. Consequentialism is flexible in the second sense but not so much in the first. Deontological approaches, by contrast, are not very flexible in either sense, given that moral rightness was determined at T1. The Reparability Principle does not settle moral rightness at T1. If the risk imposed at T1 is no longer reparable at T2, then it is no longer permissible. Thus, it may be permissible to hunt some animal at T1, but not after T2, when they have become endangered. The Reparability Principle is thus less of a decisive rule and more of a measuring stick used throughout and by which we make sure that the potential harm does not exceed the potential reparability over time. When the Reparability Principle is so applied, it is very likely that we would detect any divergence from that proportionality very early on, and that any damages beyond what could be repaired would be relatively small. If they are small enough, they could perhaps still be made reparable by allocating more resources, etc., to that end. This means that, should the risks be greater than what can be repaired, they must be reined in, or else new resources must be transferred for reparations. If the epistemic risks are significant enough, such that serious unknown

risks cannot be dismissed, this might require that alternative options need to be held open, allowing for a change of plans should the selected action turn out to entail risks that could lead to irreparable outcomes. It thus seems that the Reparability Principle is epistemically flexible in both senses: moral rightness is not fixed at T1 (before the epistemic situation changes), and given that the riskiness of an action must not exceed what can be repaired, and that reparability in part depends on our own actions, it would seem that we could change the course of action in light of new information without too much cost or difficulty.

Now, I will make a few comments on moral damage as a result of moral uncertainty. If we act on the basis of a deontological principle that later turns out to be morally mistaken, there is a high likelihood that what we did was morally wrong. The same goes for decisions made on the basis of maximising outcomes: should it turn out that we maximised the wrong value, or that maximisation was the wrong principle, then it is very likely that what we did was morally wrong. However, given that the Reparability Principle is conservative and is biased towards the status quo and reparations, the departure from what once was will never be so great as to prevent the restoration of things to their previous state. Thus, we could, in principle, change things back and apply the new principle without losing much more than time, money, and opportunities. Even so, we would have avoided irreversible moral damages and catastrophes. When it comes to risks that may affect generations in the distant future, this advantage is far from trivial.

#### WORKS CITED

- Broome, John. 2008. "The Ethics of Climate Change: Pay Now or Pay More Later?" *Scientific American* June: 69-73.
- Cohen, GA. 2012. "Rescuing Conservatism: A Defense of Existing Value." In *Finding Oneself in the Other*, 143-174. Princeton: Princeton University Press.
- Dasgupta, Partha. 2008. "Discounting Climate Change." *Journal of Risk and Uncertainty* 37/2: 141-169.

- Ferretti, Maria P. 2016. "Risk Imposition and Freedom." *Politics, Philosophy, and Economics* 15/3: 261-279.
- Finkelstein, Clare. 2003. "Is Risk a Harm?" *University of Pennsylvania Law Review* 151/3: 963-1001.
- Frick, Johann. 2015. "Contractualism and Social Risk." *Philosophy & Public Affairs* 43/3: 175-223.
- Gardiner, Stephen M. 2006. "A Core Precautionary Principle." *The Journal of Political Philosophy* 14/1: 33-60.
- Hansson, Sven O. 2004. "Philosophical Perspectives on Risk." *Techné* 8/1: 10-35.
- Hansson, Sven O. 2005. "Seven Myths of Risk." *Risk Management* 7/2: 7-17.
- Hansson, Sven O. 2007. "Philosophical Problems in Cost-Benefit Analysis." *Economics and Philosophy* 23: 163-183.
- Hansson, Sven O. 2007b. "Hypothetical Retrospection." *Ethical Theory and Moral Practice* 10/2: 145-157.
- Hansson, Sven O. 2009. "From the Casino to the Jungle: Dealing with Uncertainty in Technological Risk Management." *Synthese* 168: 423-432.
- Hayenhjelm, Madeleine and Jonathan Wolff. 2012. "The Moral Problem of Risk Impositions: A Survey of the Literature." *European Journal of Philosophy* 20: E26-E51.
- IPCC. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva. <http://ipcc.ch/report/ar5/syr/>.
- Jackson, Frank. 1991. "Decision-Theoretical Consequentialism and the Nearest and Dearest Objection." *Ethics* 101/3: 461-482.
- Kumar, Rahul. 2015. "Risking and Wronging." *Philosophy & Public Affairs* 43/1: 27-51.
- Lenman, James. 2000. "Consequentialism and Cluelessness." *Philosophy & Public Affairs* 29/4: 342-370.
- Lockhart, Ted. 2000. *Moral Uncertainty and its Consequences*. New York: Oxford University Press.
- Lunshof, Jeantine. 2015. "Regulate Gene Editing in Wild Animals." *Nature* 521/7551: 127.
- Mariscal, Carlos and Angel Petropanagos. 2016. "CRISPR as a Driving Force: The Model T of Biotechnology." *Monash Bioethical Review* 1: 1-16.
- McCarthy, David. 1997. "Rights, Explanation, and Risk." *Ethics* 107/2: 205-225.
- Nozick, Robert. 2006. *Anarchy, State, and Utopia*. Malden, MA: Blackwell.
- Oberdiek, John. 2009. "Towards a Right against Risking." *Law and Philosophy* 28/4: 367-392.
- Parfit, Derek. 1987. *Reasons and Persons*. Oxford: Clarendon.
- Pritchard, Duncan. 2015. "Risk." *Metaphilosophy* 46/3: 436-461.
- Railton, Peter. 1985. "Locke, Stock, and Peril: Natural Property Rights, Pollution, and Risk." In *To Breathe Freely: Risk, Consent, and Air*. Edited by Mary Gibson, 89-123. Maryland: Rowman & Littlefield.
- Roeser, Sabine. 2018. *Risk, Technology, and Moral Emotions*. New York: Routledge.

- Sahlin, Nils-Eric and Johannes Persson. 1994. "Epistemic Risk: The Significance of Knowing What One Does Not Know." In *Future Risks and Risk Management*. Dordrecht: Kluwer, 37-62.
- Sepielli, Andrew. 2013. "Moral Uncertainty and the Principle of Equity among Moral Theories." *Philosophy and Phenomenological Research* LXXXVI/3: 580-589.
- 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.
- Smith, Allan H., Elena O. Lingas and Mahfuzar Rahman. 2000. "Contamination of Drinking-Water by Arsenic in Bangladesh: A Public Health Emergency." *Bulletin of the World Health Organization* 78/9: 1093-1103.
- Steigleder, Klaus. 2016. "Climate Risks, Climate Economics, and the Foundations of Rights-Based Ethics." *Journal of Human Rights* 15/2: 251-271.
- Sunstein, Cass R. 2000. "Cognition and Cost-Benefit Analysis." *The Journal of Legal Studies* 29/S2: 1059-1103.
- Sunstein, Cass R. 2010. "Irreversibility." *Law, Probability, and Risk* 9/3-4: 227-245.
- Thomson, Judith Jarvis. 1986. "Imposing Risks." In *Rights, Restitution & Risk: Essays in Moral Theory*, 173-191. Cambridge, MA: Harvard University Press.
- Thomson, Judith Jarvis. 2001. *Goodness and Advice*. Princeton, NJ: Princeton University Press.
- Weisbach, David and Cass R. Sunstein. 2009. "Climate Change and Discounting the Future: A Guide for the Perplexed." *Yale Law & Policy Review* 27/2, 433-457.
- Zimmerman, Michael J. 2008. *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.

## NOTES

1. Russian roulette cases have become standard in the risk literature; one of the earliest employments of the example may be Nozick's (2006, 74).

2. Thomson distinguishes between three cases: the imposition of harm, the imposition of pure risk, and the imposition of impure risk. 'Impure risk impositions' are 'mixed' in that they consist in both the imposition of harm and the imposition of risk. Such risks can be imposed in a single act or in a series of acts. Thomson's distinction is specifically introduced in the context of a rights-based perspective but it seems to have more general implications than that. In particular, many environmental risk impositions may consist in an actual harm to the environment and a risk imposed on humans. This is not how Thomson intended it (both harm and risk of harm are cases of harm to individuals and their rights), but it brings light to something interesting.

3. What I am after here is a general model of thought, or a habit of thought, more than a precise stance, that tends to focus on these two points in time: that of the decision, on the one hand, and that of the consequences, on the other. This implicit distinction between a time when, or just before, the act is performed, and a time when all the outcomes are known, re-occurs in various ways in very common examples and distinctions, such as the *ex ante/ex post* distinction,

foreseen/actual consequences distinction, or the doctrine of double of effect, to name a few. The deontological stress on practical reason, reasons and deliberation in contrast to the outcomes also implicitly promotes such a dual-time model. Likewise, the consequentialist stress on total value of consequences, also seem to promote this way of thinking. Take, as an example, the well-discussed example in Frank Jackson's 1991 paper on the nearest and dearest objection: a doctor is facing a choice between different medicines for her patient (1991, 462ff.), one that could either kill or fully cure, and one that could only partially treat but neither fully cure nor kill the patient. This example is premised on two different times that are epistemically and morally distinct: the time of decision (when the actual effect of the risky but potentially curing drug is not known) and the time of consequence (when the actual outcome will be known).

4. See Zimmerman (2008) for an excellent discussion about whether it is the best actual outcomes, the subjectively expected best outcomes, or the prospectively expected best outcomes that ought to be the morally significant category.

5. This idea is derived from Sahlin and Perssons' distinction between epistemic risks and outcome risks (1994). They write about a "[...] completely different type of risk [to outcome-risks], seldom discussed but equally problematic and difficult to handle: the risk of not knowing or not knowing enough" (1994, 37f.). They continue: "High level decision makers are often presented with (research) reports of some kind or another, on the basis of which a decision is to be taken. These reports may well give accurate and trustworthy information about the outcome-risks considered, but what about the things not considered, and what if the scientific and intellectual quality of the reports is not that dependable? Basing a decision, it may be a high level decision or a low level decision, on scanty or inaccurate information will inevitably lead to unnecessary risk-taking – to epistemic risk taking" (1994, 38). They further argue that a "[...] lack of robust knowledge leads to a considerable epistemic risk, which in turn means that we cannot accurately monitor the outcome-risks involved" (1994, 37f.).

6. Cf. Sepielli: "What is it rational for you to do when you are uncertain between conflicting moral views? This is not the old question of what you should do, given some moral theory, when you are uncertain about the non-moral facts. The question I'm asking takes us back one step; it asks what it is rational to do when you're uncertain regarding the theories themselves" (2013, 580).

7. This section is intended to highlight two particular worries rather than to provide a comprehensive overview. There are of course many interesting theoretical approaches in the literature (see Hayenhjelm and Wolff 2011). In addition to rights and consequences, interesting recent contributions have extended the moral framework to, for instance, freedom (Ferretti 2016), contractualism (Kumar 2015; Frick 2015), and moral emotions (Roeser 2018).

8. See, for example, Lenman (2000) on the Epistemic Argument against consequentialism. On the predication problem of CBA, Hansson says the following: "Our inability to predict how different alternatives will develop into the future affects not only cost-benefit analysis but also other policy-guiding practices. It puts limits to our ability to plan rationally for the future, with cost-benefit analysis or with any other method. Arguably, however, this problem is particularly conspicuous for methods like CBA that conforms with consequentialist principles. Consequentialism refers to all consequences of an action, including consequences that occur far off in a distant

future and/or depend on very complex causal mechanisms. Therefore it is a problem for consequentialism that we cannot know the distal consequences of our actions” (2007, 171).

9. See, for example, Hansson (2004, 146) and Hansson (2007, 180f.) on compensability.

10. Nozick (2006) develops a similar idea but within a framework of rights, such that an imposition of risk on P could be justified if it is fully compensated for *post ante*. This avoids the second problem of not respecting persons: it would have to be the person whose rights had been violated that was compensated in order to justify the risk from a rights-based perspective.

11. This way of talking, of course, goes against the standard framework of cost-benefit analyses, where states of affairs are typically measured in terms of preferences, i.e. how they are valued. In that context, any talk about ‘abstracting value’ that supposedly existed in the states themselves would be a bit anathema.

12. All losses would, on monistic versions of utilitarian frameworks, be considered simple losses: all losses and gains can be traded and exchanged for each other. On the assumption of incommensurability and pluralism, this would not be the case. However, such losses could presumably be traded and compensated for with something else of commensurable value. Here, I am assuming that there are losses that are not in this way compensable. Without that assumption, reparability would collapse into cases of compensated for and not compensated for harms. This could allow for natural exploitations against monetary gains, provided the benefits of the former exceeded the costs of the latter.

13. In most cases, these two concepts seem to largely overlap: if something is irreversible, then it cannot be returned to how it was, and if something is irreparable, then it cannot be repaired such that it is returned to the state it was in. In both cases, the damage would be permanent. In some ways they differ: reparability is only relevant to states of affairs, but irreversibility is also relevant in the contexts of progressive harm over time, escalations, and decisions and policies that have a direction. I shall, therefore, prefer ‘irreparability’ when discussing states of affairs and ‘irreversibility’ when it comes to policies, decisions, and trajectories over time. The two also differ in other ways: irreparability already implies a negative impact (that would warrant a need for reparation) in a way that irreversibility does not. Furthermore, reparation is consistent with an improvement, whereas reversal seems to give little room for anything but a return to a previous state of affairs.

# *Risk-Averse Sufficiencyarianism*

## *The Imposition of Risks of Rights-Violations*

### *in the Context of Climate Change*

Lukas H. Meyer and Harald Stelzer  
University of Graz, Austria

**ABSTRACT.** Climate change induced uncertainties raise important challenges to normative theory. By their actions, people currently living do not necessarily harm future people, but rather they impose risks of rights-violations on them. This raises important questions. How should we weigh the climate change induced impositions of risks on future people? How can we comparatively assess the imposition of these risks on the interests of currently living people and the costs, harms and forgone benefits induced by climate change policies? By outlining a risk-averse interpretation of intergenerational sufficiencyarianism, we provide a framework for answering some of these questions. After developing basic elements of intergenerational sufficiencyarianism, we argue in favour of a risk-averse interpretation of weak sufficiencyarianism and discuss the implications of this approach in the context of responding to climate change.

**KEYWORDS.** Sufficiencyarianism, intergenerational justice, risk imposition, uncertainty, basic needs

#### I. CLIMATE CHANGE UNCERTAINTIES AND THE IMPOSITION OF RISKS OF HARMS

While we know that climate change implies severe risks for the basic wellbeing of future people (IPCC 2014), we lack certainty as to what the consequences of our actions and policies today will be in the future and what risks will materialize. The environmental effects of anthropogenic climate change are subject to a multitude of uncertainties due to the complex and nonlinear nature of the climate system and related aleatory variability, as well as our limited knowledge of the climate system, possible emission scenarios, the relations between emissions and temperature rise (climate sensitivity) and between temperature rise and its impacts on the environment. Additionally, there are uncertainties involved

in the relation between environmental consequences and human wellbeing. Even if we knew the environmental effects of climate change, their impacts on social, economic and political systems would still be uncertain since they depend on factors like the exposure of these systems, their vulnerability to climate change induced consequences and their resilience to climate change more generally.

We cannot say for certain that our actions will harm future people. Rather, many of our actions impose risks of harms on them. This leads to important research questions in normative theory. Can one say that the imposition of risks harms future people, and if yes, in what way (Zimmerman 2006)? Does the imposition of risks of harms constitute a rights violation (Thomson 1986; Perry 2007; Steigleder 2016)? How do we determine the magnitude of a risk of harm? How do we make trade-offs between risks of harms, considerable costs and forgone benefits?

## II. BASIC NEEDS SUFFICIENTARIANISM

Answers to these questions will depend on the normative theory employed. In the present contribution we will concentrate on intergenerational distributive justice by employing a sufficientarian conception (e.g. Meyer 2003; 2009a; Page 2006; 2007).

Sufficientarianism, as understood here, considers the protection and realization of basic interests to be of utmost importance. Sufficientarianism relies on a threshold conception that specifies how well off people ought to be in terms of the fulfilment of, for example, their (rights-) protected basic interests. According to strong sufficientarianism, benefits to people below a threshold have lexical (absolute) priority over benefits to people above the threshold (Shiffrin 1999; Meyer 2009b). Here we will argue that a weak form of sufficientarianism can be justified for assessing actions with uncertain or risky consequences. Weak sufficientarianism gives some, but not lexical priority to the benefits of people below the threshold.

Of course, there are many questions concerning how we define the sufficientarian threshold precisely. What dimensions of wellbeing need to be included? If there is more than one such dimension, how should one weigh them? Also, where exactly should we place the threshold if sufficiency of wellbeing is to be realized? As an approximation, one can use a basic needs approach, as it entails a qualitative threshold that distinguishes lives that have a certain minimum quality from lives that lack this quality (see Copp 1998, 122; Miller 2012; Gough 2015). Moreover, the specification in terms of basic needs also justifies the normative significance that sufficientarianism ascribes to its threshold of sufficiency. Basic needs are “intrinsically morally demanding” (Brock and Reader 2004). The fact that something is a basic need by itself entails that it ought to be met, and that doing so matters more than meeting non-basic needs or mere preferences (Braybrooke 1987; Brock and Reader 2004; Thomson 2005, 175; Miller 2012).<sup>1</sup>

Basic needs are often defined in an ostensive way, especially by providing a list of needs that are considered basic, such as needs for survival, food, water, or autonomy (see, for example, Streeten 1979). Nevertheless, we prefer a definition based on the notion of harm, in particular necessary and serious harm. P has a basic need for O, it is argued, if and only if P’s not having, being, realizing, etc. O harms P, and this harm occurs necessarily and is serious. According to David Copp, “[m]atters of basic needs are things anyone would require in some quantity and in some form in order to avoid a blighted or harmed life” (1998, 124; Wiggins and Dermen 1987; Doyal and Gough 1991; Thomson 2005; Miller 2012). Even though such an understanding is close to material needs, it is not restricted to them. It has been argued that the non-satisfaction of P’s basic needs harms P by impairing P’s autonomy or rational agency (Doyal and Gough 1991; Copp 1998; Page 2007); by preventing P from fully functioning in certain ‘core human activities’ such as working, having a family, or playing (Braybrooke 1987; Miller 2012); or by preventing P from flourishing (Anscombe 1958; Kowarsch and Gösele 2012).

### III. RIGHTS-BASED APPROACH

This understanding of harm in terms of basic needs can be used to interpret sufficientarianism as reflecting an interest-based rights approach. The following four claims indicate our understanding of how the notions of rights, interests, and basic needs are interrelated and why this understanding is relevant for an interpretation of intergenerational sufficientarianism. First, needs can be understood as closely related to important interests. As David Wiggins puts it: “If a person needs *x*, then he has an interest in *x*’s being or becoming available to him.” And if the need in question belongs to the basic needs, then “[...] his having *x* may be said to represent a vital interest of his” (1998, 17). Second, according to Joseph Raz’s well-known interest theory, “[a]n individual *X* has a right if and only if *X* can have rights, and other things being equal, an aspect of *X*’s well-being (his interest) is a sufficient reason for holding some other persons to be under a duty” (1986, 166). Third, if we follow a broad understanding of justice (see Mill 1863, Chapter 5), justice considerations apply to intergenerational relations if future people will have rights that have to be respected by people presently living, who in turn stand under correlative duties to future people. Fourth, we can assume a close connection between basic needs and important interests as well as the idea of people having rights reflecting the protection and realization of their important interests. If so, then allowing people to realize their basic needs and protecting people against falling below this sufficientarian threshold level of well-being is a matter of fulfilling duties of justice, since falling below this threshold implies the infringement or violation of (basic) rights of persons.<sup>2</sup> Based on a sufficiency threshold, it seems possible to spell out which rights must be fulfilled or to what degree (as we will discuss below), in order to reach a certain level of wellbeing above the sufficiency threshold.

The bearers of moral rights claims are primarily individual persons. This is important as we do not only aim to maximize overall wellbeing, but also to consider how wellbeing is distributed, taking the separateness

of persons seriously. A rights-based approach also connects neatly to the need for institutional arrangements for their protection (Caney 2007, 287). Henry Shue has argued for a social guarantee against standard threats (1980, 13), “[...] a duty to create, if they do not exist, or, if they do, to preserve effective institutions for the enjoyment of what people have rights to enjoy” (1980, 17). For Shue, such a duty does not exist against all possible violations of human rights, but only against standard threats, understood as predictable remediable threats (1980, 33). As pointed out by Bell, anthropogenic climate change can be seen as such a threat and we stand under a duty to promote effective institutions for the protection of basic human rights against the threat posed by severe consequences of climate change (2011, 111).

#### IV. CHALLENGES FOR SUFFICIENTARIANISM

As potential harms of the consequences of currently living peoples’ actions will only materialize in the future, intergenerational sufficientarianism concerns questions of risk imposition on future people (and also on currently living people when they will be among those affected). Some have argued for sufficientarianism as a general theory of distributive justice; we see sufficientarianism as a theory with considerable benefits for dealing with questions of intergenerational justice, based on a threshold conception of possible harms as a setback of the basic interests of future people.<sup>3</sup>

Against sufficientarianism it has most often been objected that it is unable to provide a substantial specification of its threshold of sufficiency (Arneson 2005; Goodin 1987; Rawls 1999; Casal 2007). First, it has been claimed that any specification of this threshold is bound to be arbitrary. Second, it has been argued that even if there is a non-arbitrary way of specifying the threshold, this threshold cannot plausibly have the immense normative significance that is ascribed to it by sufficientarianism. We indicated our (incomplete) answers to these objections in the first two

sections, above. We submit for discussion a basic-needs (intergenerational) sufficientarianism, suggesting that at least for the foreseeable future the needs of future people can be expected to be very similar to those experienced by currently living people (see Wiggins 1998; Doyal and Gough 1991; Miller 2012; Gough 2015). Accordingly, a sufficientarian threshold could be specified on the basis of the basic needs of both currently living and future people, while still allowing for differing, context-dependent basic needs of currently living and future people.<sup>4</sup> Based on this substantive threshold conception, future people have a *prima facie* rights claim against us that we do not perform actions that constitute imposing risks of harms on them by threatening the fulfilment of their important interests that reflect their basic needs.<sup>5</sup>

## V. RISK-AVERSE SUFFICIENTARIANISM

We now focus on how to interpret the imposition of risks of rights violations within a basic needs sufficientarianism and how to weigh such impositions against the certain infringement of rights. Looking at cases of risk imposition from a sufficientarian perspective we should avoid risks with respect to basic interests of wellbeing. From such a viewpoint, our focus is on the imposition of risks of rights violations in terms of people's sufficiency rights. These sub-threshold risks include the risk of people falling below or not reaching the threshold of sufficiency and the risk of people below the threshold being lowered in their wellbeing even further. If people have a *prima facie* claim against us that we do not harm them by intentionally performing actions that threaten their basic needs related to their basic interests (Zimmerman 2006; Steigleder 2016), then imposing such threats can be understood as wrongfully harming people by violating their rights.

We now turn to some more specific questions regarding what we will dub a risk-averse intergenerational sufficientarianism. By 'risk-averse' we mean a certain relation between outcomes and uncertainty. We want to

avoid outcomes that lead people to fall below the threshold or to stay there. The higher the probability of such a negative outcome, the less we want it. A higher probability of negative outcomes (defined by the threshold) counts against an option. Of course, the higher the probability of a positive outcome the better it is. That means we want policies with the least likelihood of sub-threshold risks, but we prefer a higher likelihood of an outcome if we evaluate it to be positive. In addition, and this makes the view risk-averse, we want to avoid risks as much as possible even if that means foregoing higher benefits. A risk-averse interpretation of sufficientarianism is based on an asymmetry between potential losses and potential gains with respect to the wellbeing of persons. Potential losses weigh more heavily than potential gains. If more uncertainty means that there is a larger probability of falling below the threshold, then more uncertainty counts against a policy; this is so even if more uncertainty should also mean that there is a larger chance of attaining better (or very good) outcomes.

In the case of non-probabilistic uncertainty concerning sub-threshold outcomes, we should appeal to precautionary considerations that one can understand to be risk-averse in a broader sense. We want to avoid options in which we are uncertain about their outcomes or how high the probability of certain outcomes is. Of two outcomes with the same estimate in terms of wellbeing, one should choose the one that does not exhibit uncertainty or exhibits less uncertainty.

## VI. PROBLEMS OF RISK-AVERSE SUFFICIENTARIANISM

In our understanding, a risk-averse sufficientarianism combines a rights-based approach with a consequentialist account. Several problems are nevertheless evident here, including the following. First, if we look at two policies with different levels of (probabilistic) uncertainty where neither of the two policies extends below the sufficiency threshold, the reasons for risk-aversion that we have developed so far do not apply. However,

one can augment risk-averse sufficientarianism by adding an argument for risk aversion above the threshold. The simplest way to do this is to extend the scope of the validity of the idea that potential losses weigh more heavily than potential gains. If the idea holds generally, then we have reason to avert both below and above threshold risks.

For the second and third problem much depends on how strong the priority is that we attribute to the violations of rights below the threshold. Strong sufficientarianism attributes lexical priority to the avoidance of below threshold violations of rights. If the avoidance of rights violations below the threshold is being attributed lexical priority, there would be no point in weighing these violations against rights violations above the threshold. Lexical priority means that at any time the avoidance of rights violations below the threshold is more important than other claims; under lexical priority, the number of people affected and the extent to which their other rights or benefits could be realized is irrelevant. Lexical priority thus has three important consequences: (i) as mentioned above, it prohibits weighing rights violations below the threshold against rights violations or the set back of other interests above the threshold (Birnbacher 2003); (ii) if we assumed strict priority of the avoidance of rights violations below the threshold then we would have to secure these rights for only one person or a small number of people even if that required violating other rights of many people just above the threshold; (iii) according to this understanding, actions that reduce the wellbeing of people above the threshold for the sake of improving the position of people below the threshold are not objectionable unless such reductions would bring these people below the threshold.

These implications of strong sufficientarianism may seem counterintuitive. However, even if attributing such strong priority to the rights claims below the threshold could be justified when the fulfilment and violations of rights can be understood to be certain, there are further problems when the imposition of risks of rights violations is at stake.

Thirdly, a strong interpretation of a rights-based sufficientarianism that gives the threshold lexical priority has also important implications when it comes to questions of risk imposition. Strong sufficientarianism might be understood as categorically prohibiting the pursuance of any policy that comes with any probability of causing people to fall below the threshold. Such an interpretation by itself will not allow us to choose among policies in the context of responding to climate change since. Realistically speaking, we have to assume that all sets of policies that are open to people currently living will – with high likelihood – avoid ‘dangerous climate change’ (e.g. by staying below 1.5°C compared to pre-industrial global mean temperature), which can be presumed to imply such risks. This does not preclude that, by themselves, some measures and policies which can be understood to contribute to such a set may not imply such risks. However, this clearly is a highly complex empirical issue that is outside of the scope of the present contribution. In what follows, we focus on sets of policies assuming that they come with such risks.<sup>6</sup>

Strong sufficientarianism can be understood in a more complex way that addresses this issue. We start with the claim that we ought to choose the policy with the least likelihood of causing people to fall below the threshold. Two further issues then arise. Do numbers count? And does it matter how far below the threshold? Both questions are plausibly answered positively on the basis of a priority view below the threshold. Yes, numbers count and it matters how bad people fare below the threshold. This more complex view can be stated in the following way (combining strong sufficientarianism with a negative priority view). First, the avoidance of (the imposition of the risk of) rights violations in terms of people’s sufficiency rights (and thus the avoidance of any risk of people falling below the threshold of sufficiency) has absolute or lexical priority. Second, both below and above the threshold it matters more to benefit persons the more persons are being benefited and the greater the benefits in question. But, trade-offs between persons above and below the threshold are precluded (strong sufficientarianism).

However, third, if we cannot avoid the imposition of the risk of violating people's rights as specified by the sufficientarian threshold, to avoid risking people's sufficiency rights matters more the worse off the persons would be, the more people are likely to be harmed and the greater the weighted sum of rights violations. Fourth, analogously above the threshold: to avoid risking the violation of people's other rights or risking harm to their other claims matters more the worse off the person would be, the more people are likely to be wronged and the greater the weighted sum of rights violations and harms caused.

Strong sufficientarianism can be criticized fourthly for not normatively distinguishing between the imposition of risks of rights violations and the certain violation of the same rights. In our understanding, this is a major objection that strong sufficientarianism shares with other rights-based approaches. If the protection of rights is considered absolute such that there can be no imposition of risks of rights violations, this leads to a paralysis of social life (Jackson and Smith 2006; Hansson 2013, 36; Steigleder 2016) as all or most courses of action come with the risk of the violation of some rights. Despite the fact that there are cases in which it seems obvious that imposing risks on others is impermissible, it often seems permissible or non-avoidable to impose certain risks on others, including sub-threshold risks, namely, as introduced above, the risk of bringing people below or not up to or even further below the threshold. As we already noted, if the imposition of risks of rights violations below the threshold has lexical priority, this prohibits any action that comes with such risks if we have the option of avoiding the imposition of such risks. But what if we do not have such an option? How then should we assess alternative courses of action with differing likelihoods or uncertainty of rights violations? What if the probability of sub-threshold risks is low and the benefits of these actions are high?

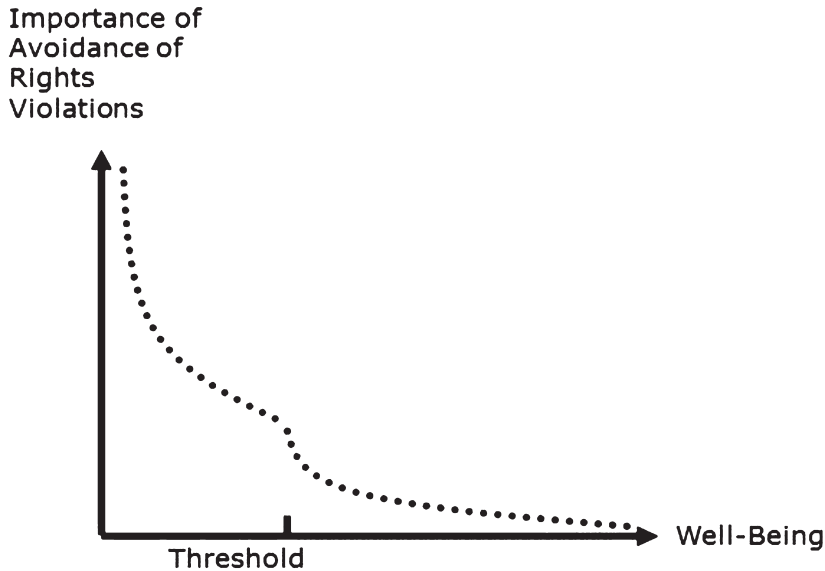
Here we rely on the answer that we sketched above in response to the third problem. The assessment of alternative courses of action, all of

which impose risks of rights violations, will rely on comparing these options in terms of (i) the interests (rights) that are under threat of being violated and the interests (rights) being fulfilled by carrying out these risk-imposing actions, namely both below and above the threshold and (ii) the number of people whose rights are being put at risk. In addition, the assessment of alternative courses of action will also rely on comparing these options in terms of (iii) the likelihood of the violation of the rights in question.<sup>7</sup>

## VII. RISK-AVERSE WEAK SUFFICIENTARIANISM

The problems of strong sufficientarianism indicate that lexical priority is not adequate, at least not for cases of risk imposition, and thus not for most cases in the intergenerational context. We therefore have to consider more complex models for assigning an order of priority of certain rights. Dieter Birnbacher (2003) distinguishes two such models for the case of weighing the protection of actual claims. The first attributes fixed or relative weight to certain rights in order to receive an order of priority. The second more complex model also takes into account the extent to which a right may be violated if after weighing up the importance of all rights the protection of another right (or other rights taken together) has priority in the case under consideration. And in our context we will want to add: even for rights, the fulfilment of which is not a matter of degree, we can be in a situation where we do not know or have only probabilistic knowledge whether the right in question will be violated or infringed and how many people will be affected. In response we should consider risk-averse weak sufficientarianism, a position that gives some, but not lexical priority to the risks of rights violations below the threshold.

Risk-averse weak sufficientarianism can be described as follows and by including most of the features of the amended version of strong sufficientarianism delineated above.

**Risk-Averse Sufficientarianism**

First, the avoidance of the imposition of the risk of rights violations in terms of people's sufficiency rights (and thus the risks of people falling below the threshold of sufficiency or of not reaching the threshold) has some priority.

Second, both below and above the threshold it matters more to avoid harm to persons the more persons are being harmed, the higher the likelihood of harms being suffered and the greater the weighted harms in question. This means that the determination of the normative weight of a risk of a rights violation depends on basic criteria. First, it matters which rights are being violated. Rights differ in their quality in the sense that their realization or non-realization will have different consequences in terms of realizing a sufficient level of wellbeing. We assume that people have moral rights whose realization is necessary for them reaching the sufficiency threshold. The realization of some of these rights will make a bigger difference in terms of avoiding sub-threshold levels of wellbeing. These effects depend on both the specific substantive understanding of

the threshold and the particular circumstances. Second, the quantity of the rights violations matters, i.e. the number of rights that will possibly be violated both below and above the threshold as well as the number of people affected by the risk of a rights violation and where they stand relative to the sufficiency threshold. This also holds for possible benefits. Both below and above the threshold it matters more to benefit persons the more persons are being benefited, the higher the likelihood of the benefits being realized and the greater the benefits in question. Such a position would allow for trade-offs between persons above and below the threshold as well as between harms and benefits. According to this view, as pointed out above, we would be able to assign special weight to risks of rights violations below the threshold. This means that even though trade-offs are possible, they would only be allowed in qualified situations and under consideration of the higher weight of risks of rights violations below the threshold and the higher weight of harms in comparison to benefits.

Third, if we cannot avoid the imposition of the risk of violating people's rights as specified by the sufficientarian threshold, avoiding risking people's sufficiency rights matters more the worse off the person would be, the more people are likely to be wronged, and the greater the weighted sum of the considered rights violations.

Fourth, above the threshold the avoidance of risking the violation of people's other rights or risking harm to their other claims matters more the worse off the person would be, the more people are likely to be wronged and the greater the weighted sum of rights violations and harms caused.

Ideally, for us to assess alternative courses of action on the basis of this position of weak sufficientarianism we would need to have knowledge about all possible consequences (harms and benefits) and their probabilities. Furthermore, those likely or possible harms would need to be understood in terms of likely or possible rights violations and their specification (see above). Additionally, we would need to assign comparative

weights to these rights that may get violated and/or infringed. Information is needed about all possibly affected persons and their position in regard to the sufficientarian threshold; given the importance of the threshold we would need to know for all possibly affected persons which potential rights violation or infringement of rights would put them below the threshold and which course of action has which potential of bringing people above the threshold.

#### VIII. IMPLICATIONS OF RISK-AVERSE WEAK SUFFICIENTARIANISM IN THE CONTEXT OF CLIMATE CHANGE

In what follows we would like to consider some implications of our version of risk-averse weak sufficientarianism in the context of climate change. This form of weak sufficientarianism belongs to the family of rights-consequentialist views that combine special weight to certain rights claims (reflecting the sufficientarian threshold) with an aggregative understanding of the claims of people both below and above the basic needs threshold and an understanding of how to weigh risks of rights violations versus certain infringements of rights. The priority of the threshold would allow us to see at least some options as likely non-starters, while at the same time being able to deal with situations in which all options come with the possibility that some people will fall under the sufficientarian threshold or further below the threshold (but in different numbers or to different degrees). Here we suggest relying on the third feature of our account of weak sufficientarianism (outlined above) for choosing the least unjust option available in such a situation. It is especially this possibility to evaluate the moral permissibility in circumstances in which all sets of policies open to people currently living may come with considerable risks of the violation of certain rights or the setback of interests of currently living as well as future people that makes our account applicable to climate change responses. Weak sufficientarianism allows us to choose between options even though they may all come with sub-threshold risks.

Furthermore, it also allows us to include interests and rights above the threshold in our considerations.

Our account is also able to include further considerations for deciding on the moral permissibility of options in the case of risk imposition. Due to the fact that it is often not possible to exclude the sub-threshold risks, we may still be able to limit these risks by the way we implement certain actions and policies. For instance, driving a car always includes a probability of sub-threshold risks for oneself as well as for others. However, we think of it as permissible to operate such vehicles due to the great benefit they promise to those who drive and to society at large. Nevertheless we can qualify the permissibility of driving a car by asking operators to act or not act in certain ways. We ask drivers to make sure that their cars fulfil certain safety requirements and we forbid people to drive under the influence of alcohol (or certain levels of alcohol in their blood). The possible influence of our actions on others implies, or so one can argue, certain duties of care. In many cases of otherwise permissible actions it is when we do not take up these duties of care that the risks imposed by these actions become morally relevant. What seems to make these actions morally questionable is that we perform them, although for all we know people may end up materially harmed by their consequences, or that we go ahead with some action without giving any consideration to the possibility that doing so could materially harm other people. From such a perspective, what we are to be blamed for – even when no harm materializes – seems to be that we (i) do not take these aspects of our actions to be conclusive normative reasons against so acting, and (ii) that we neglect alternative options available to us. One can easily see the importance of adding the idea of duties of care when it comes to the often emphasized failure of implementing more aggressive mitigation measures.

Furthermore, risk-averse weak sufficientarianism also has implications for the distribution of harms and benefits in responding to climate change, as it will require that we do not impose burdens and/or risks on

those realizing a level of wellbeing below or far below the threshold in favour of benefiting those much better off. Due to the risks involved and the potential diverse regional climate effects at least some options in the context of global climate change may not satisfy demands put forward by our position. For example, certain emission intensive activities may benefit some, while at the same time lead to the infringement of sufficiency rights of others. Many options will most likely create winners and losers, depending not only on existing climate conditions, but also on population density, economic development and the vulnerability and resilience of ecological, economic and social systems. Those geographically and economically most vulnerable to climate change and often living at (or below) the subsistence level could also be those most likely to be negatively affected by uneven effects of certain ways to respond to climate change, lacking the ability to adapt to the changing circumstances and, at the same time, being least responsible for global warming, especially considering their and their societies' historical contributions to the problem (see, for example, Meyer 2013). In principle, our responding to climate change can also have the opposite consequences: policies may benefit some of the most vulnerable and poorest countries by reducing climate change induced risks and therefore allowing them to reach a level of sufficiency or even pass that threshold.

In the case of climate change we have to consider not only the possibility of negative consequences of our actions, but also a potential violation of the autonomy of future people. Some options we choose today may limit the choices of future people. The moral significance of risking can therefore not solely rest on its perceived likelihood of, or its potential for causing material harm to the person upon whom the risk is imposed. Without doubt, risk regarded as the probability of harm is crucial for our normative assessment of climate change policies where we normally are interested not in their intrinsic impermissibility or wrongfulness, but focus rather on their possible bad outcomes and side effects. What makes these actions risky and potentially impermissible, wrong or at least

questionable, are the material harms that the consequences of these actions can ripen into. However, this account has been challenged by construing the possibility of nonmaterial harms constituted by the imposition of risks as being separate from the harms that are risked (e.g. Thomson 1986; Finkelstein 2003). For example, Oberdiek (2012) has recently developed an autonomy-based account arguing that cases of pure risk imposition can constitute a nonmaterial harm by foreclosing formerly safe possibilities available to other people.

Furthermore, missing information about outcomes in a situation of uncertainty raises additional questions. Due to issues of complexity and uncertainty, in the case of the evaluation of climate change policies we will often only be able to sketch likely outcome scenarios. The consequences of options open to people currently living will depend on things like economic development or international cooperation as well as other actions or the value of climate sensitivity. 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) (Stelzer and Schuppert 2016). In order to sketch what chances in terms of wellbeing gains and risks a certain option might offer with regard to the relevant sufficientarian threshold we need to use our best available predictive tools. However, in cases of decision-making under (non-probabilistic) uncertainty and ignorance 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.

In cases where we cannot assign probabilities, probability ranges or levels of confidence to certain outcomes, risk aversion as such does not allow us to choose among options. A question that arises in this context is whether we have to give unspecified risks of harms the same weight as risks whose probability is known. Knowledge is of course an important factor in our decision-making and possible negative outcomes we cannot assign probabilities to often speak against an option, especially if the potential negative consequences could be severe and lead people to fall

under the sufficientarian threshold. At the same time, unspecified risks can give rise to precautionary action, even if we have no proven probabilities for a risk or a clear picture of causes and effects of certain threats.

Even though risk aversion would take us in the direction of avoiding the slightest risks, especially when we may face very bad or catastrophic outcomes, we need to avoid alarmist, doomsday scenarios from paralyzing our decision-making. One way how to deal with cases in which we have some idea of their likelihood would be to include a probability threshold that prevents us from extreme risk-aversion and extreme risk-seeking. If the estimated probability of a certain negative outcome of an option would fall below a certain threshold, one reaction could be to ignore or parenthesize this outcome for our evaluation. This, however, is less than plausible in the case of potentially very bad or catastrophic outcomes. Some of these outcomes can be described as ‘fat tails’ coming with an unexpectedly thick ‘tail’ toward the edges of a distribution curve, indicating an irregularly high likelihood of catastrophic events if the low probability risk would materialize. The relevance of fat tails has been particularly emphasized in the context of climate change (Weitzman 2011).

However, as cases of mere possibilities – without a specification of the probability of future risks – are in principle unlimited, as any decision or option may have catastrophic unforeseen consequences, we have to disregard most of them.<sup>8</sup> Hansson has proposed a three-step procedure to select the consideration-worthy mere possibility arguments. The first step is to replace the mere possibility by a reference to more specified degrees of possibility (disambiguation). If this is not possible one can apply two symmetry tests. The first is a test of opposite effects. Here we investigate if the mere possibility argument “[...] can be countered with an argument showing that we have at least as strong reasons to consider the possibility of some other effect that (1) is opposite in value to the effect originally postulated [...], and (2) has equal or larger moral weight than the postulated effect” (Hansson 2013, 90). The second is the test of alternative causes. Here we investigate if a mere possibility argument

“[...] can be defeated by showing that we have at least as strong reason to consider the possibility that either the same effect or some other effect that is at least as undesirable will come about if the action under consideration is not performed” (Hansson 2013, 91). Both tests seem relevant not only for dealing with mere possibility arguments, but also for cases in which all our options will come with risks that we have strong reasons to avoid from our sufficientarian perspective developed above. One could therefore strive to include such tests in our risk-averse weak sufficientarian position.

## IX. CONCLUSION

Climate change induced uncertainties and risks raise important challenges to normative theory, especially in our relation to future people. It has been argued elsewhere that sufficientarianism can be justified in addressing particular features of the relations among non-contemporaries, such as the contingency of the composition of future people upon currently living people’s actions, the impossibility of interaction among non-contemporaries and the limits of knowing more remote future people’s more specific interests and desires (Meyer 2003; Meyer 2009a; Meyer and Roser 2009). In these previous publications, it was also shown that attributing lexical priority to a threshold defined in basic rights will make it less likely that currently living people’s duties vis-à-vis future people are over-demanding.

For the reasons given above we find that strong sufficientarianism is not adequate in addressing the problem of currently living people imposing risks of rights violations on future people. We propose responding to these challenges on the basis of what we have called a risk-averse weak sufficientarianism. We expand this view by suggesting that we should assess wellbeing both below and above the threshold in terms of a version of the priority view and by identifying several criteria for assessing the importance of the risks imposed. Such a view allows us to weigh risks of rights violations below

as well as above the threshold and against each other as well as against possible benefits of an action or a policy, while at the same time allowing for special weight to risks of rights violations below the threshold. We think that such weighting is needed as the imposition of a risk will in most cases not have the same weight as a harm that will result with certainty.

However, the assessment of the complex risk-averse sufficientarianism as sketched in the present contribution will depend on our answers to substantive questions. A central question is what rights should determine the threshold of wellbeing such that imposing the risk of violating these rights can be understood to be of particular normative relevance and, indeed, how important it is to reach the threshold or to avoid people falling below the threshold. Also, while we identified criteria for assessing risks (among them, likelihood, number of people affected, seriousness of the harm and/or normative status of the rights violation) we have not attempted to specify how one should go about assessing differing risks of rights violations, setbacks of interests as well as unrealized benefits. A fully developed version of risk-averse weak sufficientarianism would have to make important assumptions about the relative weight of (interests) rights as well as the risks of violating these rights, and would have to argue for the plausibility of these assumptions. It would further have to balance the different criteria for different cases. Even though these are complex questions and challenges for future work, answers to these questions, as shown in the latter part of our article, are required for the position of risk-averse weak sufficientarianism to guide us in choosing the least unjust option available in future situations of global environmental crises.<sup>9</sup>

#### WORKS CITED

- Alkire, Sabina. 2002/ *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*. Oxford: Oxford University Press.
- Anand Sudhir and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services." *Journal of Economic Perspectives* 7/1: 133-150.

- Arneson Richard. 2005. "Distributive Justice and Basic Capability Equality: 'Good Enough' Is Not Good Enough." In *Capabilities Equality: Basic Issues and Problems*, edited by Alexander Kaufman, 101-128. London: Routledge.
- Anscombe, Gertrude Elizabeth Margaret. 1958. "Modern Moral Philosophy." *Philosophy* 33/124: 1-19.
- Bell, Derek. 2011. "Does Anthropogenic Climate Change Violate Human Rights?" *Critical Review of International Social and Political Philosophy* 14: 99-124.
- Birnbacher Dieter. 2003. *Analytische Einführung in die Ethik*. Berlin: Walter de Gruyter.
- Braybrooke, David. 1987. *Meeting Needs*. Princeton: Princeton University Press.
- Brock, Gillian and Soran Reader. 2004. "Needs, Moral Demands and Moral Theory." *Utilitas* 16/3: 251-266.
- Caney, Simon. 2007. "Global Poverty and Human Rights: The Case for Positive Duties." In *Freedom from Poverty as a Human Right: Who Owes What to the Very Poor?* Edited by Thomas W. Pogge, 275-302. Oxford: University Press Oxford.
- Casal, Paula. 2007. "Why Sufficiency is Not Enough." *Ethics* 117/2: 296-326.
- Copp, David. 1998. "Equality, Justice, and the Basic Needs." In *Necessary Goods*. Edited by Gillian Brock, 113-133. Lanham, MD: Rowman and Littlefield.
- Doyal, Len and Ian Gough. 1991. *A Theory of Human Need*. New York: Guilford.
- Goodin, Robert E. 1987. "Egalitarianism, Fetishistic and Otherwise." *Ethics* 98: 44-49.
- Gough, Ian. 2015. "Climate Change and Sustainable Welfare: The Centrality of Human Needs." *Cambridge Journal of Economics* 39: 1191-1214.
- Hansson, Sven Ove. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. London: Palgrave Macmillan.
- IPCC. 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva: IPCC.
- Finkelstein, Claire. 2003. "Is Risk a Harm?" *University of Pennsylvania Law Review* 151: 963-100.
- Jackson, Frank and Michael Smith. 2006. "Absolutist Moral Theories and Uncertainty." *The Journal of Philosophy* 103: 267-283.
- Kavka, Gregory. 1981. "The Paradox of Future Individuals." *Philosophy & Public Affairs* 11: 93-112.
- Kowarsch, Martin and Andrea Gösele. 2012. "Triangle of Justice." In *Climate Change, Justice and Sustainability: Linking Climate and Development Policy*. Edited by Ottmar Edenhofer, Johannes Wallacher, Hermann Lotze-Campen, Michael Reder, Brigitte Knopf, and Johannes Müller, 73-90. Dordrecht: Springer.
- Meyer, Lukas H. 2003. "Past and Future: The Case for an Identity-Independent Notion of Harm." In *Rights, Culture, and the Law: Themes from the Legal and Political Philosophy of Joseph Raz*. Edited by Meyer Lukas H., Stanley L. Paulson, and Thomas W. Pogge, 143-159. Oxford: Oxford University Press.
- Meyer, Lukas H. 2009a. "Intergenerationelle Suffizienzgerechtigkeit." In *Generationengerechtigkeit. Ordnungsökonomische Konzepte*. Edited by Nils Goldschmidt, 281-322. Tübingen: Mohr Siebeck.

- Meyer, Lukas H. 2009b. "Sufficientarianism: Both International and Intergenerational?" In *Absolute Poverty and Global Justice*. Edited by Michael Schramm, 133-144. London: Ashgate.
- Meyer, Lukas, 2013. "Why Historical Emissions Should Count." *Chicago Journal of International Law* 13/2: 598-614.
- Meyer, Lukas H. and Dominic Roser. 2009. "Enough for the Future." In *Intergenerational Justice*. Edited by Axel Gosseries and Lukas H. Meyer, 219-248. Oxford: Oxford University Press.
- Meyer, Lukas H. and Sanklecha Pranay. 2016. "Philosophy of Justice: Extending Liberal Justice in Space and Time." In *Handbook of Social Justice Theory and Research*. Edited by Clara Sabbagh and Manfred Schmitt, 15-35. New York: Springer.
- Mill, John Stuart. 1863. *Utilitarianism*. London: Parker.
- Miller, David. 2007a. *Human Rights, Basic Needs, and Scarcity*. Working Paper Series. Center for the Study of Social Justice & Department of Politics and International Relations. Oxford: Oxford University.
- Miller, David. 2007b. *National Responsibility and Global Justice*. Oxford: Oxford University Press.
- Miller, David. 2012. "Grounding Human Rights." *Critical Review of International Social and Political Philosophy* 15/4: 407-427.
- Oberdiek, John. 2012. "The Moral Significance of Risking." *Legal Theory* 18/3: 339-356.
- Page, Edward A. 2006 *Climate Change, Justice and Future Generations*. Cheltenham: Edward Elgar.
- Page, Edward A. 2007. "Justice Between Generations: Investigating a Sufficientarian Approach." *Journal of Global Ethics* 3: 3-20.
- Parfit, Derek. 1986. *Reasons and Persons*. Oxford: Oxford University Press.
- Perry, Stephen. 2007. "Risk, Harm, Interests, and Rights." In *Risk: Philosophical Perspectives*. Edited by Tim Lewens, 190-209. London: Routledge.
- Rawls, John. 1999. *A Theory of Justice*. Cambridge, MA: Belknap.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon.
- Reader, Soran. 2006. "Does a Basic Needs Approach Need Capabilities?" *Journal of Political Philosophy* 14/3: 337-350.
- Schwartz, Thomas. 1978. "Obligations to Posterity." In *Obligations to Future Generations*. Edited by Richard I. Sikora and Brian Barry, 3-13. Philadelphia: Temple University Press.
- Shiffrin, Seana. 1999. "Wrongful Life, Procreative Responsibility, and the Significance of Harm." *Legal theory* 5: 117-148.
- Shue, Henry. 1980. *Basic Rights: Subsistence, Affluence, and U.S. Foreign Policy*. Princeton: Princeton University Press.
- Shue, Henry. 1993 "Subsistence Emissions and Luxury Emissions." *Law and Policy* 15: 39-59.
- Steigleder, Klaus. 2016. "Climate Risks, Climate Economics, and the Foundations of Rights-Based Risk Ethics." *Journal of Human Rights* 15: 251-271.

- 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.
- Streeten, Paul P. 1979. "Basic Needs: Premises and Promises." *Journal of Policy Modeling* 1/1: 136-146.
- Thomson, Garrett. 2005. "Fundamental Needs." *Royal Institute of Philosophy Supplement* 80/57: 175-186.
- Thomson, Judith Jarvis. 1986. *Rights, Restitution, and Risk*. Cambridge, MA: Harvard University Press.
- Weitzman, Martin L. 2011. "Fat-Tailed Uncertainty in the Economics of Catastrophic Climate Change." *Review of Environmental Economics and Policy* 5/2: 275-292.
- Wiggins, David. 1998. "Claims of Need." In *Needs, Values, Truth: Essays in the Philosophy of Value*. Edited by David Wiggins, 1-58. Oxford: Blackwell.
- Wiggins, David and Sira Dermen. 1987. "Needs, Need, Needing." *Journal of Medical Ethics* 13/2: 62-68.
- Zimmermann, Michal J. 2006. "Risk, Rights, and Restitution." *Philosophical Studies* 128: 285-311.

## NOTES

1. Against basic needs as the currency of justice it has often been objected that this theory is materialistic (Sen 1984, 1993; Anand and Ravallion 1993), and promotes paternalism (Alkire 2002). These objections are unconvincing. Contrary to the materialism objection, needs-based sufficientarianism can and typically does acknowledge basic needs for psychological or social goods (Reader 2006; Wiggins 1998). And contrary to the paternalism objection, there are various ways in which needs-based sufficientarianism can and typically does respect the autonomy of the recipients of justice. First, the theory does not demand *meeting* the basic needs of the recipients of justice, but only *enabling* them to meet these needs (Copp 1998). Often the recipients are able to decide for themselves whether or in which way to meet their basic needs. Second, needs-based sufficientarianism can and often does claim that autonomy itself is a basic need (Doyal and Gough 1991; Reader 2006; Gough 2015). And third, needs-based sufficientarianism can and sometimes does claim that autonomy is even (one of) the final end(s) of the satisfaction of any basic need (Doyal and Gough 1991; Copp 1998).

2. This is related to David Miller's idea that principles of justice reflecting general human rights can be justified with reference to people's basic needs (see, for example, Miller 2007a; 2007b, 178-200).

3. Arguably, we can develop a plausible response to the so-called 'non-identity problem' (Schwartz 1978; Kavka 1982; Parfit 1986) on the basis of adopting a threshold notion of harm (Meyer 2008; 2009; Meyer and Roser 2009). Such a notion is a central feature of (intergenerational) sufficientarianism.

4. To be sure, a fully developed understanding of the threshold based upon needs would have to specify how to weigh the non-fulfilment of basic needs in order to be able to comparatively assess outcomes that adversely affect different basic needs and more and less strongly.

5. Whether such an account can be interpreted as over-demanding is open to debate. If we take the number of future generations to be very large and if intergenerational risk-adverse sufficientarianism entails a duty to perform actions that can improve the well-being of many future generations so that they will reach levels of well-being over and above the threshold, people currently living stand under duties that could imply very significant decreases in their wellbeing (namely up to bringing them down to the level of the threshold) (Meyer and Roser 2009).

6. We would like to thank one of the anonymous reviewers for elucidating this point.

7. Hansson sees the weighing of duties or rights as a problem for deontological and rights-based theories when dealing with risk imposition. Mechanisms that draw the line between acceptable and unacceptable risks imposition by assigning decreasing weight to duties or rights as probabilities decrease are needed in order to be able to account for the associated benefits of risky actions. If we take into account this insight, this leads us away from deontological and rights based theories to a form of extended utilitarianism, “[...] in which the violation of an obligation has a negative value to be weighted against other values in a procedure that has the structure of utilitarian calculation” (Hansson 2013, 30). We agree with the thrust of this observation. Risk-averse sufficientarianism is a rights-consequentialist and, thus, aggregative view. The interesting question is which considerations should have what weight in our assessment of alternative courses of action all of which impose the risk of some rights violations.

8. “A policy to attend seriously to all such arguments would make us hostages to the constructors of fanciful scenarios. We would become victims to superstition, and our general competence as decision-makers would deteriorate” (Hansson 2013, 89).

9. We would like to thank our colleagues of the Department of Philosophy and the FWF Doctoral Program Climate Change of University of Graz for helpful discussions. Special thanks to Kian Mintz-Woo for his insightful comments on the penultimate draft and to the anonymous reviewers.

# *On the Criteria of the Rightful Imposition of Otherwise Impermissible Risks*

Klaus Steigleder  
Ruhr University Bochum, Germany

**ABSTRACT.** While it is often held that rights-based moral theories cannot deal adequately or convincingly with risks this paper is part of a project to develop a compelling rights-based risk ethics. According to rights-based moral theories it is generally prohibited to impose risks on others that involve the possibility of basic harms like death or grave or severe injuries and that cannot be sufficiently controlled. But there can be circumstances when the imposition of such risks is permitted or even required. This article tries to show when this is the case. To this end five criteria are discussed and justified. Under certain conditions, the imposition of a mere risk of a basic harm may be preferable to its secure infliction. The imposition of the risk of a basic harm on one person may be justified when only in this way a risk of an even greater harm to another person can be prevented. The imposition of a small-probability risk of grave harm on one person may be situationally justified when only in this way a high-probability risk of grave harm to another person can be prevented. The criterion of the normative inevitability of risks of basic harms pertains to risks standardly connected to certain technologies, technological systems and forms of social organization. These risks can be justified if a technology or form of organization helps to avert even greater risks to all affected. Finally, the imposition of risks of basic harm can be justified by the consensus of those affected.

**KEYWORDS.** Rights, risks, agent-risks, recipient-risks, risk imposition, possibilities

## I. INTRODUCTION

In the present contribution, I would like to advance the theory of a rights-based risk ethics delineated in a recent article (Steigleder 2016). There, I argue that the problematic tendency of rights-based moral theories to prohibit any risk impositions results from a one-sided focus on the moral rights of the *recipients* of actions. Instead, one must realize that agents and recipients possess the same rights. Just as the rights of a

recipient of the actions of others are violated by the impositions of certain risks, the rights of an agent are violated when she is not allowed to impose certain risks on others. For risks are ubiquitous and almost any action involves risk impositions. If an agent was not allowed to impose any risks on others, this would gravely restrict her ability to lead her life. Therefore, the potential effects of the prohibition of risk impositions must be weighed up with the potential effects of the imposition of these risks on others. This will allow us to delineate morally permissible and impermissible risk impositions.

It is a question of proportionality. If the effects of the materialization of a risk on the rights of a recipient are small relative to the effects the avoidance of the imposition of this risk has on the rights of an agent, then, all else being equal, it is permissible for the agent to impose the risk on others. If, on the other hand, the effects of the avoidance of the risk imposition on the rights of an agent are small relative to the effects the materialization of the imposed risk will have on a recipient, then it is impermissible for the agent to impose the risk. Thus, in the context of risk impositions, it is possible to make a basic *normative* classification of risks into *agent-risks* or *A-risks* and *recipient-risks* or *R-risks*. *A-risks* are risks that an agent is permitted to impose on others, because a prohibition of these risks would unduly privilege the rights of the recipients over the rights of agents. *R-risks* are risks that an agent is in principle prohibited from imposing on others, because the permission of these risks would unduly privilege the rights of agents over the rights of the recipients.<sup>1</sup>

In contrast to my 2016 article in which I especially focused on A-risks, in the present contribution I will focus on the most important kind of R-risks that consist of considerable possible harms to others, like death or severe injuries, and that are not sufficiently controllable by an agent. Thus, it is in principle reasonable for an agent to forego actions that impose such risks. For even though the avoidance of the imposition of R-risks will significantly restrict an agent in her freedom, such a restriction is proportionate in view of the importance of the possible harms to

the affected persons.<sup>2</sup> But while it is generally prohibited to impose R-risks on others, one may nevertheless be justified or even required to do so under certain circumstances. Our topic here can thus be designated: the development and justification of the criteria of the rightful imposition of R-risks.

Before I continue, however, some further preliminary clarifications are appropriate. First, I understand *risks* as possible negative effects or possible harms and distinguish them from *chances* as possible positive effects or possible benefits or gains. The imposition of a risk on another person means that a new or additional possibility is created by an action (or complex of actions) that the person will suffer a harm. This possibility is primarily *epistemic* in nature. Due to our limited knowledge, we have to assume that certain developments and outcomes are possible even if this is in fact not the case. The unknown liquid in a glass may be poisonous; we have to reckon with this possibility, and we would impose an R-risk on a child if we let the child drink from it. This will even be the case if the liquid is harmless, because we do not know this. Thus, I suggest distinguishing *realistic possibilities* as possibilities we have reason to assume due to our epistemic situation from *real possibilities* (e.g. the decay of a certain radioactive isotope during the next 24 hours) which exist in the event that something is objectively open or undetermined. Real possibilities are undetermined; realistic possibilities can even exist under determined circumstances.<sup>3</sup> Risks are primarily epistemic in nature because they primarily involve realistic possibilities.

Second, I use *risk* as a generic term comprising both risks in a technical sense, where an attempt is made to measure the possibility of outcomes with the help of probability numbers and to treat the possibility as more or less likely, and uncertainty, where probability numbers are not or cannot be given. Actually, when speaking about risks, what I mean at least most commonly is uncertainties (regarding possible negative outcomes).

Third, I will presuppose that all persons possess an equal and cardinal normative importance and for this reason an entitlement to be able to

lead their lives. Thus, they possess equal moral claim rights to the conditions necessary for being able to lead their lives, such as life itself, physical and psychological integrity, to name only the most basic features.<sup>4</sup> These rights are both negative and positive rights, i.e. rights to non-interference and to the assistance of others. Positive rights or rights to assistance are always situationally conditioned. Persons have a right to be helped when they are not able to help themselves and another person is able to help at no comparable cost (see Gewirth 1978, 218f.; 1996, 42f.).<sup>5</sup> Often, such help can only be provided at no comparable cost when it is provided collectively by way of institutionalizing the help or by supporting the respective (welfare state) institutions (see Steigleder 2014, 473f.). However, the institutionalized fulfilment of positive rights does not remove their conditionality, even if the criterion of ‘no comparable cost’ has to be adapted. Thus, for instance, the right to life as the most basic necessary condition of being able to lead one’s life does not automatically comprise, for instance, of (positive) rights to very expensive medical treatment or life-sustaining machinery.

As the criterion of ‘no comparable cost’ presupposes that the moral claim rights of persons form a hierarchy according to their indispensability for being able to lead one’s life.<sup>6</sup> For the necessary conditions for being able to lead one’s life are ‘necessary’ conditions insofar as they are all *on the whole* indispensable. But they are nevertheless indispensable to different degrees, ranging from being absolutely indispensable – such as life itself – to conditions whose absence is more or less life-threatening, and then further spanning from conditions whose absence makes it (almost) impossible to *lead* one’s life to conditions whose absence severely restricts one’s ability to lead one’s life and finally to conditions that can, to different degrees, be situationally or temporarily dispensed with.

Situationally, therefore, a less basic right of one person may be superseded by a more basic right of another person. For the temporary transgression of a right of one person can be justified if the effect of this transgression on the ability of the right-holder to lead her life is a minor

one in comparison to the massive threat to this ability of the other person, which is averted by the transgression. Think of Joel Feinberg's example of a hiker who is surprised in the mountains by dangerous weather conditions and breaks into an untenanted cabin in order to protect herself. While the life or health of the hiker and therefore her ability to lead her life is acutely threatened, nothing comparable is to be expected to be the case for the owner as a consequence of the damage done to and of the use of his cabin (in Feinberg's example, the hiker not only breaks into the cabin but also burns part of its furniture in a stove [1980, 230]).

Finally, I presuppose that the rights of persons include a right to the effective protection of their rights. This makes necessary certain institutions like the state, but also international and global institutions. Thus, the justification of the already mentioned institutions of the welfare state is part of a more general justification of rights-protecting institutions.

## II. CREATING RISKS AND RESPONDING TO EXISTING RISK CONSTELLATIONS

The simplest case of a risk imposition is the creation of a new or additional risk for one or more persons through a single actor's action (or sequence of actions). But risks can also be created by the implementation of complex technologies, by the cumulative effects of the actions of many people, by collective actions, by the implementation or alteration of policies and by institutional or structural change. In these (more) complex cases, it is notoriously difficult to make sure that wrongful risk impositions are recognized in advance and avoided. However, it is to be expected that in all such cases the criteria for assessing which risk impositions are justified and which are not will in principle be the same. For they all share the 'imposition' paradigm, i.e. the notion of the *creation of a new or additional* possibility of harm. With this the internal perspective of the (potential) participants or risk creators is prevalent. Those who are not involved in the potential risk creation must take on the perspective of the potential contributors to the risk creation to be able to evaluate their contributions.

For this evaluation must ultimately be based on the comparison of the different effects of the permission and prohibition of the contributions on both agents and recipients. The imposition paradigm becomes stretched, of course, when it is applied to institutional or structural changes that we perceive more as the result of anonymous forces being operative behind the back of agents than as the result of the imputable actions of these agents. Besides, the strategy to compare the effects on agents and recipients is confronted with special difficulties when in the context of cumulative effects the individual contribution seems to ‘make no difference’.

The perspective of risk creation must be supplemented, however, by a different perspective, namely the perspective of *existing* risks or risk constellations (e.g. the existing infection risks in hospitals, the risk of accident on certain roads or in connection with current traffic regulations). The normative question here is: who (e.g. individuals, groups, governments) ought to *respond* in which ways to existing risks? These risks will often be the results of past actions, but can also result from natural forces or events or they can be the combined product of natural forces and human actions. Since many different risk constellations usually exist simultaneously, the relevant question to ask is often the following: the elimination or reduction of (the risks of) which risk constellations should receive priority and which risk constellations may or must be tolerated, at least for the time being? In the face of existing risk constellations, the external perspective is prevalent, for the potential agents are often not those who have caused the risk constellations and often do not have special obligations to those affected by different risk constellations.<sup>7</sup> Therefore, in the face of different risk constellations a maximizing strategy can be the right prioritization criterion, i.e. it can be justified to focus on those strategies that promise to prevent the most fatalities and thus, for instance, to concentrate more on road than on rail safety (Wolff 2006, 413f.; 2007). We might also be required to focus on the most effective means, e.g. on those measures that promise to prevent more fatalities for

the same or lower costs than other measures. The fact that these strategies of maximization or efficiency can be justified in the context of the handling of existing risk constellations does not mean that they may automatically be transferred to the context of risk impositions.<sup>8</sup> Further, in the context of risk constellations these criteria may also be situationally outweighed by other criteria, e.g. the prioritization of the risks to especially vulnerable or disadvantaged people.

Strategies to reduce or eliminate existing risks will usually involve the creation of new risks. Thus, the imposition of risks often cannot be evaluated in isolation, but must be evaluated as an attempt to reduce existing risks. This must be taken into account in the following investigations, the focus of which will be the imposition of risks.

### III. WHY THE IMPOSITION OF R-RISKS CAN BE RIGHTFUL: A CRITERIA OVERVIEW

Let us return to the imposition of R-risks, i.e. of risks of basic harms that cannot be sufficiently controlled. Why does the prohibition against imposing R-risks hold only *in general* so that there can be exceptions from the rule and sufficient justifying reasons for the imposition of R-risks? The answer to this question lies both in the nature of risks and in the nature of rights, especially in their ranking or hierarchical order. Risk impositions can violate the rights of others because they create the realistic possibility of depriving another person, temporarily or permanently, of something she needs for being able to lead her life. A person has rights to the conditions necessary for leading her life. Thus, she is entitled to not be deprived by others of these conditions. As the imposition of the risk creates a realistic possibility of depriving her of conditions she is entitled to, her entitlements include the right against all other persons to refrain *as far as possible* from creating a realistic possibility of such a deprivation.

To this one might want to object that a realistic possibility is an epistemic notion and that one can therefore ‘create’ a realistic possibility

without creating an objective danger of depriving a person of conditions she is entitled to. The liquid in a glass may be harmless, a pistol may not be loaded, and so on. However, we have no other viewpoint available than our limited epistemic perspective. Thus, we must act on the basis of what we have reasons to hold possible. And a realistic possibility does not mean a mere logical possibility, but the presence of indications that objective dangers *may* be connected with certain actions. The liquid in a glass may indeed be toxic, and a pistol may indeed be loaded. We have to reckon with such possibilities. Because we as agents must evaluate the possible consequences of our actions *ex ante* in light of our epistemic situation it must be normatively relevant to us that we must deem it realistically possible that our action will infringe on the rights of others. And due to their rights, the potential recipients of our actions must hold that we are obliged to refrain as far as possible from creating the realistic possibility of infringing their rights.

The ‘possibility’ in these normative evaluations has both a normative and a ‘factual’ meaning. An agent may not be able to avoid certain risk impositions without a disproportionate restriction of his own rights. This is the rationale behind *A-risks*. But it may also not be possible, in the specific situation, for the time being or more generally, to prevent or reduce serious R-risks without measures or actions that involve ‘less serious’ R-risks. The comparison of R-risks can, on the one hand, relate to the quality of the possible *effects*. In this connection, for instance, the risk of the infringement of a more important right can be a more serious risk than the risk of the infringement of a less important right. The comparison of R-risks can, on the other hand, relate to the possibility or likelihood of the effects. In this perspective, the more probable R-risk can be more serious than the less probable R-risk of the qualitatively same effect. Evaluations of the seriousness of R-risks become difficult when these two aspects are mixed in a reversed way. How, for example, does the seriousness of a highly probable R-risk of a qualitatively lesser effect compare to an improbable R-risk of a qualitatively more important effect?

The interesting question is whether or how far the possibility and likelihood of effects extends or changes the standard criteria of ‘deterministic’<sup>9</sup> rights-based ethics. As already pointed out, one such criterion is that situationally a more important right of one person can override the less important right of another person. Starting from this criterion of the situational priority of rights two possible ways of the justification of the imposition of R-risks are initially conceivable. On the one hand, the risk imposition could be justified by *deontic inheritance* (i.e. by the justification of the deontically weaker proposition by its entailment in a justified deontically stronger proposition).<sup>10</sup> If in certain situations even the direct infringement of a right is justified when only in this way a more important right can be protected, then it must also be justified to impose in similar situations the risk of the infringement of the same kind of right if already by this (and nothing short of this) the more important right can be protected. For instance, if it is justified (obligatory) to kill an aggressor in order to protect innocent people, it is also justified to sedate the aggressor with a life-threatening sedative provided this is sufficient for the protection of the innocent people.<sup>11</sup> On the other hand, the criterion of the priority of rights suggests the existence of a criteriological analogue on the level of risks, namely that the imposition of a (qualitatively) smaller R-risk (e.g. the risk of a severe injury) on one person in order to prevent the imposition of a greater R-risk (e.g. the risk of death) on another person could be one of the criteria of the rightful imposition of R-risks we are looking for. One may call this criterion *the situational priority of preventing the risk of greater harm*.

But can it also be justified to impose a low-probability risk of severe harm, like death or severe injuries, on one or more persons in order to prevent or reduce the high-probability risk of severe harm for another person? If this can be justified, there will be an additional criterion of the rightful imposition of R-risks that may be called *the situational priority of averting the high-probability risk of grave harm*.

A further criterion may be called *the normative inevitability of R-risks*. R-risks can be normatively acceptable if they contribute to the prevention

of greater risks for all affected and if the greater risks can only be prevented by the imposition of the R-risks in question (or of comparable R-risks). A final and quite different criterion to be considered here is *consensus*. The consensus of those affected *can* make the imposition of R-risks acceptable.

#### IV. A CLOSER LOOK AT THE CRITERIA

##### *Deontic Inheritance and the Possibility of the Avoidance of the Certain Infringement of a Condition of Agency*

As observed above, it may be permissible or even required to infringe on the less important right of one person in order to protect the more important right of another person. If this protection can also be achieved in other situations by not directly infringing the right, but by merely imposing the risk of the infringement of the right in question on the other person, then, all else being equal, such risk imposition will also be permissible. And if in the same situation both possibilities, the certain infringement and the merely possible infringement of the right, are open to the agent in order to protect the more important right, then, all else being equal, the agent ought to choose the risk imposition in preference to the certain infringement of the right. For this retains the possibility that the direct infringement of a necessary condition for being able to lead one's life can be avoided. This criterion of the *possibility of the avoidance of the certain infringement of a necessary condition of agency* can come in different variants and applications. I suggest treating it less as an independent criterion, but as incorporating an important justifying reason that is relevant for the justification of different criteria.

Here, the original justification of the risk imposition is by *deontic inheritance* and a case of the justification of the weaker proposition by its entailment in a justified stronger proposition. However, in the context of deontic claims (as well as other modal propositions) particular circumspection is required. The weaker deontic claim is only entailed in the stronger if it

partakes in the justification of the stronger claim and does not introduce or is not connected with the introduction of new morally relevant facts. If, for instance, it is situationally justified to kill an aggressor in order to save innocent people, then to impose (merely) the risk of death on the aggressor will only partake in the original justification if the risk imposition also ensures that the innocent people will be saved. Now, it is doubtful that this can be done by a ‘pure risk imposition’ (to use the terminology introduced by Judith Thomson [1986, 173-176]), i.e. without physically harming the aggressor, say, by shooting him in his stomach or his leg instead of right in his heart. But in such cases of ‘impure risk imposition’, where the risk imposed is a consequence of the infliction of a direct physical harm, the moral evaluation of the risk is not independent from the moral evaluation of the harm. This is not to say that the infliction of the harm is unjustifiable, but that in cases of ‘impure risk imposition’ the justification of the risk imposition cannot simply be done by deontic inheritance. However, this does not mean that the criterion of deontic inheritance is inevitably useless or superfluous. For deontic inheritance may enable us to recognize that the imposition of a grave risk is justified by seeing that one could even be justified to directly inflict the harm connected with the risk in question if necessary.

One may be tempted to rename the criterion of the *possibility of the avoidance of the certain infringement of a necessary condition of agency* (which is itself a shorthand for “possibility of the avoidance of the certain infringement of a necessary condition for being able to lead one’s life”) into *possibility of the avoidance of the certain infringement of a right*. But while this formulation is perhaps a little more elegant, it would be misleading. To see why this is the case can help us gain a better understanding of the justifying reason behind the criterion.

Persons have both rights not to be directly infringed on their necessary conditions for being able to lead their life and not to be exposed to (particular) risks of being infringed on their necessary conditions for being able to lead their life. Thus, specific risk impositions will constitute an

infringement of the rights of the affected people and insofar as this is the case the infringement of the right will be no less certain than the infringement of a right that consists in directly infringing on a necessary condition to lead one's life. The crucial point is not the difference in the certainty between the respective infringements, but in their relative importance. Insofar as the wrong of the risk imposition consists in the creation of a realistic possibility of inflicting a specific harm, the direct infliction of the harm is, all else being equal, the morally more serious action because the (mere) risk imposition not only involves the realistic possibility of inflicting the harm, but also the realistic possibility that the harm will not be inflicted. Hence, in direct conflict the risk imposition may be the morally preferable option to the certain infliction of harm. While this may be trivial when the certain harm and the potential harm affect the same person, the difference in seriousness may also be relevant when different persons are affected.

It is important to note that the difference in seriousness is relevant in cases of a direct conflict between the certain and the only possible infliction of the qualitatively same harm or within a certain qualitative spectrum of different harms. However, it is not the case that a risk imposition is always to be preferred to the direct infliction of harm. Instead, in comparison to the creation of grave or catastrophic risk, the direct infliction of a harm may be the morally preferable or justified option. Furthermore, when no conflict is given the right not to be exposed to R-risks is as strict a right as the right that one is not directly harmed in certain ways. Finally, as the rights of persons are based on their normative importance, persons can also be wronged by actions that betray disrespect of this normative importance. In this connection, even the imposition of otherwise trivial risks can violate the rights of persons.<sup>12</sup>

### *The Situational Priority of Preventing the Risk of Greater Harm*

As already observed, the criterion of the situational priority of preventing the risk of greater harm suggests itself as a criteriological analogue

on the level of risks to the criterion of the situational priority of rights. But a problem of the justification of the risk ethical criterion is that the risks involved are R-risks and the R-risks which are here to the fore are risks of serious harm that cannot be sufficiently controlled. Thus, one is not normally morally permitted to impose *such* risks on others. One question then is whether it can be situationally permissible to impose a qualitatively less serious R-risk *of this kind* in order to avoid the imposition of a qualitatively more serious R-risk *of the same kind* on another person. This can only be justified by appealing to the criterion of the situational priority of rights, if this criterion itself justifies the direct infringement of a less important *basic* right of one (innocent) person in order to protect a more important *basic* right of another person.

Feinberg's example does not fit here because the infringement of property rights is not a basic harm (if the property is not needed to avert basic harm) and a risk of damaging the (non-necessary) property of another person constitutes an A-risk *insofar* as the damage can (almost) fully be compensated and the agent who imposes the risk is able and willing to compensate for the damage in the event that it materializes. If he is not able to compensate, the risk will be an R-risk but of another kind than the R-risks I am primarily concerned with in this paper. However, the right to transgress a property right in order to save oneself or another person is independent of one's ability to compensate the owner. Thus, all else being equal, one is justified to impose the risk of uncompensated damage to property in order to prevent, for example, the risk of death or major injury to oneself or others, and we can know this from the criterion of the situational priority of rights. Think of a trailer rolling downhill that could hit a group of children playing below. You cannot stop the trailer, but you are able to change its direction with the risk that it will now hit an uninhabited cabin. You are permitted and, if you can act at no comparable cost, even required to do so independent of your or others' ability to compensate the owner of the cabin.

However, the criterion of the situational priority of rights also pertains to the level of basic harms and rights. It may be right to injure a person in a severe but comparatively minor way in order to avert grave injuries or death to another person. Correspondingly, it is right to impose the risk of severe but comparatively minor injuries in order to avert severe injuries or death to another person. Such cases confront us with the problem of a non-arbitrary ordering of injuries. But this does not constitute an insurmountable obstacle as long as we are able to assign clear-cut cases to our morally relevant categorizations. The relevant categories are (i) minor injuries, (ii) severe injuries, (iii) grave injuries. There are indeed clear-cut cases that fit into one of these categories. A scratch normally constitutes a minor injury, a broken arm or a ruptured eardrum are examples of a severe injury, paraplegia or the loss of sight are examples of grave injuries. The rationale behind the categorization is informed by the varying impact of injuries on the ability of a person to lead her life. One criterion of the distinction between severe and grave injuries is whether the respective impairment of the ability to act or to lead one's life is transitory or permanent. But such criteria need both to be refined (e.g. long-lasting but not permanent) and supplemented (e.g. intensity of pain). Thus, considerable work has to be done here. But it is warranted to assume that in principle injuries can be evaluated in a morally relevant way and that in addition to this the proposed categories can be extended to other basic harms besides direct physical injuries such as psychological harms or the effects of environmental impacts on people's ability to lead their life. However, the lines between the categories will inevitably be fluid. Again and again, there will be cases in which it is not clear whether an injury is still minor or already severe or still serious or already grave. And it is a task of normative ethics to determine how such unclear cases are to be dealt with, e.g. on which side one ought to err or whether it is justified to fix the cases on the level of a society by way of a common understanding or of a democratic decision.

To return to the question of risk, the imposition of the risk of minor injuries normally constitutes the imposition of an A-risk, while the impositions of the risk of severe or of the risk of grave injuries both constitute the imposition of an R-risk. While the imposition of such R-risks is normally impermissible, it can situationally be right to impose the risk of a less important basic harm on one (or more) person(s) in order to avoid the imposition of a more important basic harm on one or more other person(s). Thus, the criterion of the situational priority of preventing the risk of greater harm holds and it is justified by the situational priority of rights.

*The Situational Priority of Averting a High-Probability Risk of Grave Harm*

Thus far, when comparing R-risks the focus was on the different quality of the possible harms involved. But with regard to the criterion of the situational priority of averting a high-probability risk of grave harm the perspective is a different one. For the focus is now on the “quantitative” difference between R-risks, i.e. on differences of the probabilities of the qualitatively same basic harms.

In order to gain a better understanding of the proposed criterion and to test it let us start with the following problem or question: is one permitted, and if so under what circumstances, to avert an existing risk of basic harm to one person by imposing a lower probability risk of the same basic harm on another person? And to keep things manageable, let us restrict the risks in question first to risks of grave harm. The range of possible affirmative answers to the question can be demarcated by two quite different answers. On the one hand, one might hold that whenever an (existing) risk of grave harm to one person can be averted by the imposition of a lower risk of the same grave harm on another person one is permitted to do so. On the other hand, one might hold that only when a high probability risk of grave harm to one person can be averted by imposing a very low probability risk of the same harm on another person one is permitted to do so.

Consider the following four scenarios as variations of the trailer example. Again, a trailer is heading downhill toward a group of playing children. Again, the trailer cannot be stopped, but its direction can be changed toward a cabin. But we now add the following information. The group of children in question are four in number and they are playing on a playground with a slide and other equipment. The trailer is loaded with highly explosive material. So, whether it hits the playground or the cabin, it is highly probable that the material in the trailer will explode.

*Scenario I:* The cabin is inhabited by the Millers, parents and two children, but there is a chance that the Millers are not in the cabin at the time the trailer would hit it.

*Scenario II:* The Millers are in the cabin. It is not improbable, but not highly probable that the rolling trailer will hit the playground, but the probability that the trailer will hit the cabin is a bit lower if its direction is changed than the probability that the trailer will hit the playground if no one intervenes. (Let us assume, somewhat unrealistically, that nobody will be harmed if the rolling trailer hits neither the playground nor the cabin.)

*Scenario III:* Here the same background conditions as in scenario I hold true (so with a very high probability that the trailer will either hit the playground or the cabin), but in the new scenario one can be almost sure that the Millers are not in the cabin although one cannot completely exclude that the Millers are in the cabin.

*Scenario IV:* The same conditions as in scenario I hold true, except that it is only probable (but not highly) that without intervention the trailer will hit the playground. It is still highly improbable that the Millers are in the cabin at the time the trailer would hit it.

I am especially interested in the scenarios I and III and would like to argue that while one is permitted (and, if one can do it at no comparable cost, even required) to redirect the course of the trailer in scenario III, this is not the case in scenario I. The reason for this is that only Scenario III differs significantly (enough) from the cases and criteria with which ‘deterministic’ rights-based ethics is concerned. Deterministic rights-based ethics (convincingly, I suppose) holds that normally one

may not sacrifice one person (or a minority of persons) in order to save another person (or a majority of persons). I will take this standard or principle here as justified. Now, the (well-founded) rationale behind averting the course of the trailer in scenario III is not that the Millers should die instead of the playing children, but that this offers an alternative for action which, on the one hand, will save people from grave harm and, on the other hand, goes with the justified expectation that no one will be (gravely) harmed.

One may want to object here that this course of action seems to be justified only as a result of the way in which it is framed. The purported rationale becomes more dubious when it is framed in a different, but nevertheless adequate way, namely as saving people from grave harm by hazarding the potentially grave harm of other people. But this objection is not appropriate if there is only a very low chance (in the sense that it can almost completely be excluded) that the other people (in our scenario the Millers) will be affected. In such a situation, it would instead betray an inadequate concern for the rights of those which can almost certainly be saved. But, of course, this rationale is not applicable and the objection does hold in the case of averting the course of the trailer under the conditions of scenario I.

Thus, it is important that there is indeed only a very low probability that other people will be affected and that accordingly the estimation of the probability is well-founded. I cannot go into detail at this juncture, but suffice to say that the estimation need not be expressed in probability numbers and that I would like to avoid indicating probability numbers that can be considered ‘very low’. Instead, I would like to confine myself to indicating that estimations of probabilities as ‘very low’ can indeed be well-founded. For instance, concerning the Millers it may be the case that one knows that the Millers inhabit their cabin only during their holidays in August and that it is now May, that they do not rent it out to others, and that besides, in the event of a very unlikely exception this year, one would have seen the Millers or at least heard of them.

One must, of course, examine further how strict the condition of a very low probability is and whether, and if so how much, it can be relaxed. (I assume that it is pretty strong, but that it is unavoidably vague if we do not relate it to probability numbers and that the remaining vagueness marks the justified leeway concerning the condition.) A further question to be examined is how ‘high’ the probability of the harm to be averted must be. Here I suppose that the permissible probability range can be much wider than with the other probability condition, but that the probability of the harm to be averted has to be considerably higher than the very low probability that the ‘other people’ will be affected. Thus, it may be justified to change the course of the trailer in scenario IV, but it is definitely not justified to do so in scenario II. Furthermore, we must ask what range of harms is allowed. There seems to be no need to confine the criterion under consideration to grave harm. The criterion is justified if the qualitatively same basic harms are involved for the people to be saved and the ‘other people’. This implies that the criterion is also justified if more serious basic harm can be averted by the imposition of a very low probability risk of less serious basic harm on other people. But is it also permitted to avert a high probability risk of less serious harm by means of the imposition of a very low probability risk of a more serious basic harm (grave harm) on other people? I doubt this, but the question also needs further examination.

A final problem to be discussed is what variations in the numbers of the affected people is allowed. I purposely arranged the scenarios such that the people potentially staying in the cabin are not fewer than the children playing on the playground in order to avoid a constellation that is similar to a ‘Trolley Problem’. However, the criterion will also be applicable to situations in which the number of the ‘other people’ is higher than the number of the people from which the high risk of basic harm is to be averted. The question to be further investigated is how much higher the number may be. There must be a limit, otherwise the risk imposed would turn into a low probability catastrophic risk the imposition of which is normally impermissible.

So, I conclude that the criterion of the situational priority of averting a high-probability risk of basic harm is a justified criterion. Further examination may show that it even holds under somewhat more relaxed conditions.

### *The Normative Inevitability of R-Risks*

The above criteria bear on certain pressing conflict situations in which it is hoped that certain risks of severe or grave harm to one person can be avoided by either imposing the risk of smaller harm or a quantitatively (much) smaller risk of the same harm on another person. As distinct from this, the criterion of the *normative inevitability of R-risks* justifies practices and institutions that standardly impose R-risks. These risk impositions are justified if they contribute to the prevention of greater risks for all affected and if the greater risks can only be prevented by the imposition of the R-risks in question (or of comparable R-risks). The criterion justifies the use of certain technologies, technological systems and certain forms of social and economic organization.

The ‘ability’ to prevent the greater R-risks that come with the practices or institutions in question can change with time and differ between various societies according to their different levels of development. Thus, lower safety standards may, within certain margins, be temporarily justified in a less developed country because of its more limited resources. The risks connected with certain technologies may at some point turn out to be no longer tolerable due to certain technological improvements. The condition that the greater R-risks can be avoided *for all* affected will often only be fulfilled if certain insurance schemes are established in connection with the implementation of technologies or if such insurance schemes are a constitutive part of the social or economic organization itself.

Take as an example a market economy. I dare to claim that given certain framework conditions like the welfare state a market economy is at least for the time being justified by the criterion of the normative inevitability of R-risks. To be sure, a market economy is connected with

many R-risks – different kinds of systemic risks like a breakdown of the functioning of the economy itself, for instance due to the risks linked with financial markets, or risks like environmental disasters due to the unsustainability of the existing market economies or certain existential risks for individual participants as the downside of competition. However, given certain framework conditions that incorporate certain insurance schemes, a market economy is able to avert even greater R-risks for all members of a society like starvation, poor sanitation, bad health or inadequate health care. At least for the time being, the existing unsustainability of the market economies must be tackled *within* the system as we do not know of a better alternative to a market system nor do we know how to manage the transition to the alternative system in the time required to avert the risks of, for example, catastrophic climate change in a morally defensible way.

Financial markets are indispensable for the sustainable functioning of developed market economies. However, financial markets are inevitably connected with systemic risks. Now, if a market economy is justified by its contribution to the safeguarding and protection of the rights of the members of a society, and if the financial markets are for their part required for the functioning of a developed market economy, then the systemic R-risks inevitably linked to the financial markets are justified as well. But this justification of systemic financial market risks is instrumental and hence conditional. The R-risks associated with the financial markets are justified by the criterion of the normative inevitability of R-risks *only insofar* as the financial markets contribute to the sustainable functioning of a market economy and are indispensable for it. As the current financial system does not fulfil this condition, the system and the many systemic risks it involves are largely unjustifiable. We can see from this that, far from being simply affirmative, the criterion of the normative inevitability of R-risks can function as a critical measure.

The criterion of the normative inevitability of R-risks is a far-reaching criterion and is possibly one of the most important criteria of risk ethics.

However, it is difficult to handle and it can easily be misused. Thus, it should be specified for different areas of application. On the one hand, one must try to clarify what it means for risks to be normatively inevitable (or not) in different contexts, for instance, in the context of financial markets or in the context of the use of technical instruments. On the other hand, one should try to develop sub-criteria for the evaluation of the risks standardly imposed in the different contexts.

The equality of the rights of people excludes that R-risks are standardly or systematically imposed on some people or some group of people without the compensating benefit of an increase in the protection of their rights through averting greater R-risks. Thus, there must be such a benefit *for all* people (i.e. for each and any person) on which the R-risks in question are standardly or systematically imposed. But one has to be aware that differences in the protection of the rights of the affected may result from their different abilities to protect themselves from or adapt themselves to the risks imposed on them. For the same reason, the benefits from the aversion of the greater R-risks may be more or less pronounced for different people. But here it may well be the case that the less advantaged will profit more or more directly because they do not possess additional means for averting or successfully dealing with the 'greater' R-risks. If this is the case, it would allow for the argument that the differences in the weight of the R-risks imposed may be tolerable at least to some extent. Besides, far from constituting by itself an objection to the criterion of the normative inevitability of the imposition of R-risks, the differences in the weight of the imposed R-risks call for the introduction of *accompanying* protections, especially for the most vulnerable.

This brings me back to the issue of insurance. Society has to function as an insurance for all its members, providing guaranteed assistance to those who will be affected by the materialization of the risks imposed. This is justified because (*ex hypothesi*) all benefit from the implementation of risk-laden technologies or social practices, while the risks themselves will materialize unevenly and will then lead to damage only for individual

persons or certain groups of persons. The corresponding guarantees allow for the avoidance of existential uncertainties, thus protecting the rights of the affected and especially of the most vulnerable parties. The members of the given society must contribute to such insurance schemes differently according to their different abilities.

### *Consensus*

The consensus criterion means that we are often permitted to impose R-risks on others if the affected people give their explicit, free and informed consent to the risk imposition.<sup>13</sup> While the criterion of ‘free and informed consent’ is a core criterion of contemporary medical ethics and has thus triggered a considerable body of work on its requirements and problems (see, for example, Faden, King and Beauchamp 1986; Beauchamp and Childress 2013, Chapter 4), I would like to confine myself here to highlight two problems that are especially pertinent in the context of risks. On the one hand, as comprehensive psychological research has shown, people are prone to multiple delusions and fallacies in connection with the understanding and assessment of risks (see, for example, Kahneman, Slovic and Tversky 1982; Fishhoff 1985; Kahneman and Tversky 2000; Slovic 2000). Thus, it is both difficult and necessary to ensure that the consent is based on an adequate comprehension of the risks involved. As the pertinence of the consensus criterion can make an otherwise impermissible risk imposition permissible, the task to make sure that the apparent consensus is indeed based on an adequate comprehension of the risk falls mainly on the person who wants to impose the risk. It is the task of risk ethics to contribute to the clarification of the criteria both of an ‘adequate comprehension of risks’ and of how one can make sufficiently sure that the criteria are fulfilled.

On the other hand, not every rational decision meets the freedom requirement of the consensus criterion. It might be rational for a potential employee to prefer working under the influence of dangerous chemicals to

being unemployed if accepting these working conditions is the only means to provide for her family or herself. But this does not imply that the consensus criterion is fulfilled and that the employer is permitted to expose the worker to the chemicals, and this is especially not the case if effective protective measures can be adopted relatively easily. However, many intricate problems are involved here which are known, for example, from the debates about ‘sweatshops’ and other areas of severe exploitation.

However, even if reasonable safety measures have been taken there will remain certain important tasks that can only be performed under considerable risks. We can indeed think of situations in which a person makes the informed and voluntary choice to accept such risks – in exchange for sufficient monetary compensation. Under such circumstances, given the rights of the employee and the employee’s family, the employer has the duty to assume certain responsibilities and to make sure that the employee and his family receive sufficient support should the risks materialize. If the performance of such dangerous tasks is in the public interest (think, for example, of the work of a bomb disposal squad), to ensure that such measures are taken is (also) a societal responsibility.

One of the most intricate problems of normative ethics is completeness, and I do not claim that the criteria of the rightful imposition of R-risks I have delineated and tried to justify here make up a complete or exhaustive list. However, I hope that the criteria treated here will prove useful for the advancement of a rights-based risk ethics.<sup>14</sup>

## WORKS CITED

- Beauchamp, Tom L. and James F. Childress. 2013. *Principles of Biomedical Ethics*. New York: Oxford University Press.
- Faden, Ruth F., Nancy M.P. King and Tom L. Beauchamp. 1986. *The History and Theory of Informed Consent*. New York: Oxford University Press.
- Feinberg, Joel. 1980. *Rights, Justice, and the Bounds of Liberty. Essays in Social Philosophy*. Princeton: Princeton University Press.

- Fischhoff, Baruch. 1985. "Cognitive and Institutional Barriers to 'Informed Consent'". In *To Breathe Freely. Risk, Consent, and Air*. Edited by Mary Gibson, 169-185. Totowa, NJ: Rowman & Allanheld.
- Gewirth, Alan. 1978. *Reason and Morality*. Chicago: University of Chicago Press.
- Gewirth, Alan. 1996. *The Community of Rights*. Chicago: University of Chicago Press.
- Hansson, Sven Ove. 2003. "Ethical Criteria of Risk Acceptance." *Erkenntnis* 59: 291-309.
- Hansson, Sven Ove. 2013. *The Ethics of Risk. Ethical Analysis in an Uncertain World*. London: Palgrave Macmillan.
- Kahneman, Daniel; Paul Slovic and Amos Tversky (eds.). 1982. *Judgment Under Uncertainty. Heuristics and Biases*. New York: Cambridge University Press.
- Kahneman, Daniel and Amos Tversky (eds.). 2000. *Choices, Values, and Frames*. New York: Cambridge University Press.
- MacLean, Douglas. 1986. "Risk and Consent: Philosophical Issues for Centralized Decisions." In *Values at Risk*. Edited by Douglas MacLean, 17-30. Totowa, NJ: Rowman & Allanheld.
- MacLean, Douglas. 2012. "Ethics and Risk." In *Handbook of Risk Theory* Vol. 2. Edited by Sabine Roeser, Rafaela Hillerbrand, Per Sandin, and Martin Peterson, 791-804. Dordrecht: Springer.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic.
- Perry, Stephen R. 2003. Harm, History, and Counterfactuals. *San Diego Law Review* 40: 1283-1313.
- Peterson, Martin; Christian Seidel. 2013. Unpublished Manuscript on Risk and Deontic Property Inheritance.
- Slovic, Paul. 2000. *The Perception of Risk*. London: Earthscan.
- Steigleder, Klaus. 1999. *Grundlegung der normativen Ethik. Der Ansatz von Alan Gewirth*. Freiburg: Alber Verlag.
- Steigleder, Klaus. 2014. "Human Dignity and Social Welfare." In *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives*. Edited by Marcus Düwell, Jens Braarvig, Roger Brownsword and Dietmar Mieth, 471-476. Cambridge: Cambridge University Press.
- Steigleder, Klaus. 2016. "Climate risks, climate economics, and the foundations of rights-base risk ethics." *Journal of Human Rights* 15: 251-271.
- Steigleder, Klaus. 2018. "Climate Economics and Future Generations." In *Towards the Ethics of a Green Future: The Theory and Practice of Human Rights for Future People*. Edited by Marcus Düwell, Gerhard Bos and Naomi van Steenbergen, 131-153. London: Routledge.
- Thomson, Judith Jarvis. 1986. *Rights, Restitution, and Risk: Essays in Moral Theory*. Cambridge, MA: Harvard University Press.
- Wolff, Jonathan. 2006. "Risk, Fear, Blame, Shame and the Regulation of Public Safety." *Economics and Philosophy* 22: 409-427.
- Wolff, Jonathan. 2007. "What is the Value of Preventing a Fatality?" In *Risk. Philosophical Perspectives*. Edited by Tim Lewens, 54-67. London: Routledge.

## NOTES

1. These characterizations are taken from Steigleder (2016, 264).
2. This characterization is taken from Steigleder (2016, 265).
3. This characterization is taken from Steigleder (2016, 261).
4. There are, of course, different theories on why people possess equal moral claim rights, what these rights consist in and how the rights of persons can be justified. As it is not possible to defend here the theory I presuppose (see, for example, Gewirth 1978; Steigleder 1999), I must confine myself to giving a rough outline of some of its essential features.
5. This directly follows from the equality of rights. If I was required to help you even if you could help yourself or, say, to save your life by sacrificing my own life, then, all else being equal, your rights would be more important than my rights.
6. One can only hold that people possess equal claim rights and that these rights are also positive if one also holds that the claim rights of people form a hierarchy. Thus, theorists who deny a hierarchy of rights also deny the existence of (original, non-contractual) positive rights. A further consequence of the denial of a hierarchy of rights is that a convincing rights-based risk ethics would be impossible as can be seen from the discussion of risks in Nozick (1974, 73-78).
7. A government may rightfully prioritize certain risk constellations over others due to the special obligations it has to its citizens, but may not have special obligations to the different citizens affected by the selected risk constellations.
8. This is for instance done in the cost-benefit analyzes of mainstream climate economics. See Steigleder (2018).
9. See Hansson: "Throughout the history of moral philosophy, moral theorizing has referred predominantly to a deterministic world in which the morally relevant properties of human actions are both well-determined and knowable. [...] Ethics still lives in a Newtonian world" (2003, 291). See also Hansson (2013, 1).
10. My views on this topic were formed by my engagement with an unpublished paper by Martin Peterson and Christian Seidel on Risk and Deontic Property Inheritance, on two different versions of which I commented in 2013 and 2014.
11. The example is taken from Martin Peterson and Christian Seidel's unpublished paper, but I evaluate and treat it differently in the present contribution.
12. For these reasons, I consider Stephen Perry's labeling of rights not to be exposed to certain risks of harm as 'second order' or 'secondary' rights to be unhelpful, if not misleading. See Perry (2003, 1306).
13. I hold that only in exceptional cases like an acute emergency in which an explicit consent cannot be given the assumed consent of an affected person may be sufficient for the rightful imposition of R-risks. For an instructive treatment of different kinds of consent in the context of risk impositions see MacLean (1986; 2012, 796-800).
14. I would like to thank an anonymous reviewer for helpful comments and suggestions to improve this article.

# *Risk and the Libertarian Case for Redistribution*

Jeffrey Haines  
Florida State University, US

**ABSTRACT.** In the present contribution I argue three things. First, political libertarians should think that it is wrong to create a risk of non-negligible harm for non-consenting others. Second, that in order to secure consent to being placed at risk, we will have to make compensatory payments to those we put at risk. Third, since such a system of payments would be highly impractical, and would require an enormously intrusive bureaucracy to enforce, libertarians should instead prefer a generalized system of redistribution. Such a system would base receipt of payments not on need, but on exposure to risk, and would require greater payments not from the wealthier, *per se*, but from those who create more risk for other people. In general, however, to the extent that wealthier people create more risk for the poor than vice versa (e.g. the driver of the car creates a bigger risk for the pedestrian than the pedestrian does for the driver of the car), such redistributive payments will match up with more standard notions of welfare. If this is right, then considerations of risk imply that libertarians should support a fairly robust redistributive state.

**KEYWORDS.** Risk, redistribution, libertarianism, compensation

## I. INTRODUCTION

A common principle of libertarianism that there exists not only a duty not to wrongfully harm others (typically couched in terms of not violating their negative rights) but also a rectificatory duty to compensate for wrongful harms (hereafter harms) done to others (Nozick 2013, 152-153). While simple and intuitive, this principle has well-known complications when it is applied to the continuing effects of historical harms. It also, I will argue here, has unexamined complications when it comes to exposing others to risk.

Harm is itself a complex notion, and any duty to not harm is far from being a “very simple principle” (Mill 1978, 9). Does free market competition

harm the losers of that competition? Does offending someone harm them? For the purposes of this paper, I will bracket these and other questions that Mill and others have grappled with and assume that we agree upon a notion of harm. Certainly, while there are some cases that are unclear whether they should count as harms or not, there are many others cases where the answer is clear cut.

If we accept that it is wrong to harm people (keeping in mind that by harm I mean an action that we agree is both wrong and harmful), it seems that we also need to accept a corollary risk principle that holds something like this:

If it is wrong to knowingly cause some foreseeable (wrongful) harm to a person, it is wrong to knowingly cause some non-negligible risk of this same foreseeable (wrongful) harm to that person.

If it is wrong for me to, say, knowingly shoot Jones in the arm, it is wrong for me to knowingly create some non-negligible risk that I will shoot Jones in the arm. Thus if it is wrong for me, with say perfect knowledge of the future, to pick up a gun, aim it at Jones's arm, and fire, knowing that I will shoot him in the arm, it is also wrong for me to pick up a gun that I am not sure is loaded, aim it at Jones's arm, and pull the trigger. I know that in this second case I cause a non-negligible risk of a wrongful harm to Jones.

There are two important caveats to point out about this example. First, we do not need to hold that it can never be acceptable to create risk of a harm for somebody else, just as we do not in fact hold that it can never be acceptable to harm somebody else. The prohibition on harm can be overridden by other considerations. If shooting Jones in the arm is the only way, somehow, to prevent everybody in the world from dying, I arguably should shoot Jones in the arm.

Second, we might be tempted to think that in this example the wrongness of aiming the gun at Jones is due to our *intention* to cause Jones harm (or else, why aim a possibly loaded gun at him?), but this is incorrect. An

intent to cause harm to others is not a necessary condition for making a harmful or risky activity wrong. I might aim the gun at an apple on his head from a distance and pull the trigger in order to impress him with my sharpshooting skills, and any resulting harm was not intended. This is gross negligence, but more mundane negligence is also wrong. If I fail to properly check for oncoming traffic before going through a stop sign, I create a non-negligible risk of harm for any passengers in my car and other drivers on the road, even though I certainly do not intend that anybody get hurt. Often when we create risk for others, we do not do so intentionally.

So, as with harm, we can say that the risk principle allows that a duty to not create risk for others can be overridden, but not that an intent to cause harm is always necessary in order for the act to be wrong. Once we accept that the risk principle is a justifiable principle, it has important implications for rectificatory justice. We are inclined to accept that causing harm tends to lead to some duty to compensate. In what follows I argue that libertarianism (as well as other theories that hold that there can be duties to compensate) should hold that we ought to compensate those we place at risk.

Rectificatory duties only apply once someone has been wronged. Because of this, my discussion of rectificatory duties and risk will assume, in each case, that someone has already been wronged. This is not to suggest that wronging others is permissible, or that provided compensation is paid it may be permissible to harm others or put them at risk. In the case of Jones, above, it clearly remains wrong for me to point the gun at him and pull the trigger even if I afterwards compensate him (or his family!).

In discussing harm and risk, then, compensation alone cannot offer a fully adequate account of moral duties. It will remain the case, as with Jones and the gun, that the act which is being compensated probably should not have been performed in the first place. Alternatively, the act might be one that, all things considered, was right to do, yet still wronged someone, and that wrong requires compensation. In this case, even if the act was, all things considered, justified, this should not suggest that

compensation is the primary or the only duty that the person performing the act has. Significantly, even when an act might be, at that point in time and all things considered, justified, it might well be the case that the agent ought to have invested time and effort in finding an alternative that would do less harm to the victim or put the victim at less risk.

A duty to reduce the harm (or risk) that our actions create for others is arguably, like the duty I have to not aim the gun at Jones and pull the trigger, more important than the duty to compensate. This duty to reduce the foreseeable harm or risk that an action (or policy) will create for others itself likely gives rise to a libertarian duty to support various regulatory agencies, agencies that it has been thought libertarianism should actively oppose and that libertarians in practice have often opposed.<sup>1</sup> To use one example, around 35,000 people die in automobile accidents in the US every year (National Highway Traffic Safety Administration 2018). While some of these are drivers who put themselves at risk by driving, many others are people who were killed by drivers who exposed them to risk (in 2016 around 7,000 pedestrians and bicyclists were killed in traffic accidents). This creates, I would argue, at the least, a clear libertarian justification for stricter traffic and car safety regulations.

Despite harm reduction arguably being of primary importance, the matter of rectificatory duties is also of philosophical interest, and for the purposes of this paper I make that my primary focus. This means that I will be considering cases in which others have already been harmed or put at risk. This is not to minimize the importance of not wronging others in the first place (or of the importance of trying to reduce the amount of harm or risk we do create), but it is necessary for the purposes of keeping the focus on rectificatory duties.

## II. COMPENSATION FOR RISK EXPOSURE

We might think that only actually causing a harm can lead to a duty to compensate. After all, if no harm has been done, there is nothing to be

‘made whole’. The person put at risk has suffered no loss that must be righted. It may have been wrong to put them at risk in the first place, but they are owed nothing for merely having been placed in danger if nothing comes from that danger. Acknowledging the risk principle may require revising our intuitions about the moral acceptability of many daily activities that create risk for others, but it does not require compensating those we put at risk.

This response is intuitive but, I think, wrong. There are at least two types of reasons why rectificatory justice should require compensation for being put at risk. First, there are instrumental reasons which hold that requiring risk to be compensated will tend to produce more optimal levels of risk in society. Second, there are intrinsic reasons, which hold that providing compensation for risk is the only way risk can be justified on deontological grounds.

### *Instrumental Reasons*

When a society requires only that compensation be paid when harm occurs, it allows its members to profit by creating risk for other members and incentivizes wrongdoing. Risk-creating behaviour is, in such circumstances, incentivized. If I can gain \$100 by creating a 1/10 risk of doing \$200 worth of harm to you, and I am only required to compensate you \$200 if I actually harm you, the expected value to me of creating the risk for you is \$80. I might, then, create the risk for you, which, per the risk principle, is wrong. When putting others at risk is not sufficiently penalized, there will be incentives to place others at risk when doing so is to our own benefit, even if we require that actual *harm* be compensated.

These bad incentives, however, might be able to be corrected if the risk-creator is penalized an amount to make the risk-creating unattractive. In the above case, if I have to pay \$1,000 if I actually cause harm, I no longer can expect to benefit from creating a risk of harm to you.

This solution, however, runs into two problems. It requires that it always be possible for the harm to be sufficiently compensated, and it requires that it be possible to causally trace harms to their sources. Considering both of these issues provides two further instrumental reasons to favour compensation for being put at risk rather than merely compensation for being harmed.

Harm may be irreparable in many cases, but the case of death offers the clearest example since someone who is killed cannot be compensated. Nor is it sufficient to compensate the families of those killed since this fails to compensate the person who was most grievously harmed, it is likely that no amount of money can ‘make whole’ the loss of a loved one, and some who are killed may have no families that can be compensated.

Compensating harm also requires that causal pathways, some very complicated, be traced. Often this is difficult or even impossible, while in principle figuring out which people are put at risk is much easier. It may be impossible to determine if a given case of cancer is a result of environmental pollution from a nearby factory, for example, while it may be much clearer that the operation of the factory increases the risk of cancer in the local population. If only actual harm is compensated for, the difficulty of proving actual harm means that it should be expected to be compensated for less often than rectificatory justice requires.

### *Intrinsic Reasons*

Requiring that we pay compensation for putting others at risk should, then, be expected to result in fewer cases where people do put others at risk. Per the risk principle, this means that such compensation will result in less wrongdoing. In addition to this, there is reason to think that there is an intrinsic reason why risk, and not merely harm, should be compensated. We should think that compensating someone for exposing them to risk is required to obtain their rational consent to being exposed to the risk.

Per the risk principle, when we expose others to risk, we are doing something wrong. There are two obvious potential circumstances for when it might be acceptable to expose another to risk. The first is when the expected benefits outweigh the expected harms of the risk. The second is when the person has consented to being exposed to risk.

The first route for justifying risk is open to utilitarians but not, it seems, to liberals, libertarians, or other deontologists. The mere fact that the expected benefits of creating a risk to someone outweigh the expected harms no more justifies creating the risk than it would in creating a certain harm. Libertarians hold that it is wrong to harm others in order to benefit oneself or third parties, even if such harm is, on utilitarian calculations, less than the benefits. Liberals, similarly, should be expected to respect a separateness of persons consideration, which holds that harms to one person cannot be justified by greater benefits to another (Rawls 1999, 164).

One distinction between a risk and a harm is that a person who is put at risk could actually end up better off as a result of the activity that put them at risk, but a person who is harmed, depending on our theory of harm, might by definition have to be made worse off. Thus it might be possible for putting others at risk to be justified if they are expected to benefit more than they lose from being put at risk, while such a justification may, when it comes to doing harm, be nonsensical.

It appears, however, that those committed to autonomy, as libertarians are, should be unwilling to accept this as a justification of exposing other to risk. We can think here of Robert Nozick's example of the community music programme, which broadcasts classical music throughout the community. Even though you benefit from this programme (you tune in to their broadcasts every evening), members of the music programme cannot compel payment from you on the ground that you benefited from their programme (Nozick 2013, Ch. 5). Libertarians it seems, should broadly accept this argument. There are many potential government programmes that may benefit me if they were enacted yet this does not seem

to make it justifiable for the government to compel my support of or involvement in them. To take one example: the government decides to ban the sale of sugary drinks. While I may like sugary drinks I am better off with the tax or ban because, as a result, I am healthier. Even though I can be expected to benefit from the policy, it seems unlikely that libertarians will think that this benefit is sufficient to compel my compliance. Here is a second example: the government requires additional tax revenue to enforce the ban. Even with this additional tax, I still benefit from the ban. Even though I can be expected to benefit from the policy, it seems unlikely that libertarians will think that this benefit is sufficient to justify the government compelling me to pay to enforce it.

Non-libertarians of course may not see anything wrong with either the ban or the tax to enforce it. For them, the government might well be justified in compelling me (and taxing me) in order to make possible projects that I will benefit from. My point, however, is that it is unlikely that libertarians can take this stance. The policies in the previous paragraph are precisely the sort of policies that libertarians do advocate against and that libertarian theory tends to oppose.

If libertarians accept this argument, my argument about risk is, if anything, on even stronger grounds. If, somehow, the radio programme creates a risk of bodily harm or property damage to you, you surely should not be compelled to be exposed to this risk, even if, in net, you benefit from the programme. The mere fact that those put at risk should be expected to benefit, then, will fail to provide a libertarian justification for risk.

There may, however, be cases in which the benefits are so great that even libertarianism prescribes imposing risk on others without their consent. An example, urged by an anonymous reviewer, is that of the introduction of a new technology. The invention, production, and widespread adoption of the automobile, for example, might have created benefits so great as to justify, even on libertarian grounds, imposing risk on people without their consent. Even if, however, the overall benefits of automobile

use were very high, it is certainly not the case that everyone benefited from the new technology. Those killed or seriously injured by automobiles would not be net beneficiaries of the automobile, and on libertarian grounds it would not have been justifiable to expose them (as well as those not actually killed or injured) to such risk.

If and when expected gains are simply so great that even libertarians think it justifiable to expose people to serious risks without their consent, this would ground a duty to compensate those harmed or put at risk without their consent. In such cases, where the risk exposure is, all things considered, justified, rectificatory justice seems especially important. Similar thinking applies, after all, in most other cases where the consent of a minority is overridden for greater social gain. We might take, for example, the case of eminent domain. This is something that libertarians typically oppose, and the greater social benefits of confiscating land are not, on libertarian grounds, typically thought to justify requiring the consent of those whose land is taken. In rare cases where libertarianism might condone eminent domain because the overall benefits are so great (or it is necessary to avoid some horrible loss), there is a clear libertarian case for rectificatory justice. Thus, even if sometimes libertarianism could appeal to the great benefits imposing risk on others without their consent would bring, they still should accept that there then exists a rectificatory duty towards those who were placed risk.

Since appealing to benefits to justify creating risk for others seems to be closed off to the libertarian, I turn now to the second route. This requires that exposure to risk be consented to. Consent is a promising path for the libertarian here, since consent seems to be able to justify causing harm to another. I cannot, for example, whip another for my own pleasure without their consent, but with their consent this becomes, in libertarianism and most other theories of ethics, acceptable.

When it comes to risk, it seems that similar reasoning applies. If I skydive or bungee jump, for example, I will likely be asked to sign a waiver giving my explicit consent to being placed at risk of serious injury

or death and absolving the operator of liability, provided she has exercised the due caution her industry calls for.

Risk can also be implicitly consented to. When I choose to ride in a car with a friend, I implicitly consent to being exposed to some reasonable set of risks. I accept being exposed to the risk that her brakes might fail or her tire might blow, killing us both, for example. To be sure, I do not agree to being exposed to any risk: if she starts to drive like a maniac putting me at a level of risk that I did not implicitly consent to, she is acting wrongly.

Most of the risk that we create for others, however, has not been consented to either explicitly or implicitly. When we emit pollution that creates risk for others or when we use electricity and so cause some risk of fire for our neighbours, we do not have either their explicit or implicit consent. Risk that is neither explicitly nor implicitly consented to is a serious problem for libertarianism and other deontological theories.

It seems that we cannot help but create such risk for others yet at the same time it will often be very onerous or even impossible to obtain explicit consent from those placed at risk or to be able to reasonably infer implicit consent. We would be unlikely to consent to being exposed to risk if we would only be given compensation once we were actually harmed for the simple reason that many forms of harm are irreparable. Thus, while I might consent to you exposing me to a one in ten-thousand risk of death if you pay me \$10,000, I would not consent to being killed for \$100 million. While payments may of course not always be this high, they should be expected to be greater than zero, since at zero no individual would rationally consent to being exposed to any level of risk.

Another possibility is that, instead of requiring compensatory payments, we agree to expose each other to roughly equivalent levels of risk for the sake of a greater benefit to both of us. If we both consent to being exposed to risk that the other person creates, we both may well end up, in expectation, better off. Hence, since it is in our rational interest to live in a society where we create risk for each other, we consent to being

exposed to such risk. Here, although we do not explicitly or implicitly consent to being put at risk, we might think that we hypothetically consent, in the same way that we might be thought to consent hypothetically to being coerced by the state. Part of the social contract is an agreement to be exposed to certain risks.

Sven Ove Hansson holds a view similar to this, concluding that creating risk for others is justified when

- (i) this exposure is part of a persistently justice-seeking social practice of risk-taking that works to her advantage and which she de facto accepts by making use of its advantages, and (ii) she has as much influence over her risk-exposure as every similarly risk-exposed person can have without loss of the social benefits that justify the risk-exposure (2013, 108).

There are, however, some important limitations to this as a means to justify the risk that occurs in our own world. As Hansson notes, “[...] a practice of risk exchanges presumes at least approximate agreement among those concerned about how serious a risk is and how valuable its associated benefits are” (2013, 109). This will often not occur in practice, since people often vary widely in their assessments of risks. Consider, for example, the major disagreements about the risks and benefits of nuclear power, airport security measures, climate change, genetically modified food, and self-driving cars. In practice, permitting some degree of risk will often require exposing people to a risk to which they vehemently do not consent.

There are three reasons why libertarians in particular should have difficulty with using Hansson’s solution to argue that much of the risk we expose others to is justified and hence not requiring compensation. First it is helpful to think back to the modification of Nozick’s example of the community music programme. This programme (somehow) creates some non-negligible risk that you will be harmed or your property damaged. If you are entitled to refuse to pay for the programme (which

I assume the libertarian must hold) even if your payment is necessary for the programme to run, then you should also be entitled to refuse to be exposed to the risk that having this programme creates, even if your consent is necessary for the programme to run. This is true even if it is clear that the benefits outweigh the risk.

The libertarian must conclude that the community cannot impose this risk on you even if, per (i) exposure to the risk is part of a social<sup>2</sup> practice of risk-taking that works to your advantage and you make<sup>3</sup> use of that advantage. In other words just because you do tune in to the programme and the programme can be expected to be to your benefit, you neither can be made to pay for it nor be made to bear any risk that it creates.

Second, Hansson's principle leaves open the possibility of a principled objector, who agrees to forgo the benefit in order to not be exposed to the risk. Thus, even if we allow that accepting a benefit entails that you agree to pay for it or entails that you agree to be exposed to some level of risk as a result, the libertarian must allow the permissibility of refusing to accept the benefit. Clearly, if I do not want to be exposed to the risk of skydiving, I can choose not to skydive and forgo the benefits of skydiving. It seems, similarly, that if I do not want to be exposed to air pollution or the risk of being hit by a car, and so on, I should be able to withdraw my consent. This will give every person a veto over being exposed to risk, which gets us back to the initial challenge of risk: we will be unable to avoid creating risk for people who explicitly do not consent to being exposed to that risk without radically changing our lifestyles.

Might the onus be on those who do not want to be exposed to the risk to leave the society that exposes them to risk? It is difficult to see how a libertarian could accept this conclusion. It would be as if in the case of the community music programme you were told that if you remained in the community you must pay the fee, but you are, of course, free to leave the community. We might think as well about a case where your neighbour regularly has bonfires on dry, windy nights. If you ask

him to stop because of the risk he is creating for you, his response that you are free to move away if you do not want to be exposed to the risk is hardly adequate. Putting the burden on the person who merely wants to not be harmed, taxed, or put at risk is not something the libertarian should agree to.

The third difficulty points not to a tension between Hansson's principle and libertarianism, but with the fact that even Hansson's principle will not be able to justify many cases of risk exposure in our world. If Hansson is right, then risk that is a part of a system that works towards the advantage of those put at risk and over which those put at risk have a say is justified. Yet much risk (and benefit) will often be asymmetric. A factory farm that pollutes a river may create enormous risk for those living downstream, yet exclusively benefit its employees, management, and shareholders, who live upstream. People in rich countries emit much more greenhouse gases than do those in arid Sub-Saharan African countries or low-lying South Asian countries, and this activity largely benefits the former while putting the latter at risk. To be sure, there are actual harms, and not just risks, associated with climate change. That these harms are asymmetric is a result of the underlying risk being asymmetric. If the risk is systematically asymmetric, then Hansson's principle will not justify it being imposed on people. This means that even if libertarians could somehow make use of Hansson's principle, and I argued above it seems to be in tension with basic libertarian commitments, they will still be faced with many actual cases of risk that will require compensation.

Risk and benefits, moreover, should tend to be systematically asymmetric. That is, people with more power or wealth will generally tend to be more likely to expose those with less power or wealth to risk than vice versa. While the people in the US may create significant risk for people in Bangladesh or the Maldives, it is not the case that the people in those countries also create significant risk for people living in the US.

When risk and benefits are systematically asymmetric, there cannot be said to be a fair exchange of risk. The exposure to risk must, for

libertarians, occur only with the exposed person's consent. There is no rational reason why they should agree to be exposed to risk with no accompanying benefits. This means, that compensatory payments for being exposed to risk will likely be necessary in order to secure their consent. Without such compensatory payments, there is no way to make risk and benefits sufficiently symmetrical for consent to be likely.

Note that these compensatory payments are not payments to compensate for a wrong already committed, they are to secure, in advance, the consent of those who will be put at risk. Securing consent from those exposed to risk is one way in which we might avoid, on libertarian grounds, doing wrong by exposing them to risk.

Requiring that only harm, and not risk, be compensated has been argued to be inadequate for several reasons. Practically, it will tend to incentivize creating risk for other people. It will also be unable to account adequately for irreparable harms or cases where causally tracing harms is difficult or impossible. More importantly, it is likely that compensatory payments are necessary in order to get people to rationally agree to be exposed to risk. We might think that some social contract that involves equal exchanges of risk would work, but this leads to the problem that libertarianism should allow people to refuse to be exposed to risks that they do not accept. When risk is asymmetric, moreover, compensatory payments will be necessary in order to secure consent.

A libertarian who accepts the risk principle is in the tough position of either having to create no non-negligible risk for others or find a way to get their consent for the risk she is creating for them. Compensation is likely to be a necessary condition on being able to get or infer consent. Thus, so long as we accept that libertarianism (and for that matter many other systems of deontic ethics) ought to regard placing others at non-negligible risk without their consent as a wrong, there will tend to be a duty to compensate those placed at risk either *ex ante* in order to be able to reasonably get or infer their consent to being put at risk or, if the risk exposure has already occurred, *ex post* as required by rectificatory justice.

### III. REDISTRIBUTION

If there is a libertarian duty to compensate those who are exposed to risk, then the libertarian will, if they are part of a complex society, have to be committed to making a large number of compensatory payments. We create non-negligible risk for others every day. When we emit any pollution, we create the risk of that pollution harming someone. When we use electricity, we create a risk of fire for our neighbours. When we drive, we risk hitting pedestrians, even if we drive with meticulous care. So long as the risk we create is non-negligible, we must, on libertarian grounds, get the consent of those who are placed at risk. The large number of people killed by fires, automobile accidents, and so forth indicates that these risks are in fact far from negligible.

This need to compensate others for putting them at risk means that, in principle, libertarians should be committed to a highly complex system of payments from those creating risk to those put at risk. Such a system, in order to be workable, would require sizable institutions able to monitor risk and enforce payments. It is clear that, unless the libertarian can provide grounds for rejecting the risk principle and the need to provide compensation for risk exposure, the night-watchman state desired by libertarians would have to be a fairly complicated state, with a complex institutional scheme for enforcing compensatory payments to those harmed and those put at risk.

If being a member of a large society means that one creates risk for others, and if creating risk for others requires that one compensate them, then members of large societies will tend, in principle, to owe each other a very complex arrangement of compensatory payments. It would be wrong to assume that these payments will simply balance out, given that, as argued above, many risks are systematically asymmetric. Since monitoring the risk we create for each other and enforcing the payment of compensation would itself require a complex institution or set of institutions, additional taxes would have to be levied to fund these institutions.

A perfect level of actual measurement of risk and enforcement of compensation payments would likely be nightmarishly complex: government institutions would, for example, have to record how many people I drove past on my way to the store, determine how much risk I created for each, and then bill me the requisite amount and distribute it to them. In practice, the complex network of compensatory payments would have to be simplified and streamlined. To this end, what institutions our libertarian utopia established would have to make simplifying assumptions and rough calculations. Behaviour, such as jay-walking or drunk driving, that tends to create high risks for others might be made illegal and/or fined, with the fine either being given to those who were clearly placed at risk or, if this is not determinable, functioning as a kind of Pigovian tax<sup>4</sup> or deterrent.

Perhaps those who drove more would be required to make more payments than those who drove less or those without cars on the reasonable assumption that driving more creates more risk for others. Those who emit more pollution also might be fined/taxed on similar grounds. Indeed, any activity that created a non-negligible risk of the kind of harm that a libertarian would oppose would, given the risk principle, be subject to a compensatory taxation.

This, however, is just one end of the problem. At the other end, our libertarian compensatory institution(s) will have to determine who is owed compensatory payments. This will likely be a more difficult problem to solve than figuring out who commits the risk-generating activities. Indeed, it is the very complexity of figuring out both ends of the problem that means that the compensatory solution to the problem of risk requires an institutional approach. No individual would be likely to even be able to figure out who has put her at risk, especially once complex systems like the climate or an ecosystem are involved.

Here our institution(s) could rely on simplifying assumptions again. Those who live closer to polluting factories, for example, should be expected to be at greater risk of harm from air pollution. Those in

low-lying countries are likely at greater risk from climate change, pedestrians at greater risk than drivers, and so forth.

If it is the case that some groups are systematically placed at risk more than others, they may be eligible for higher levels of compensatory payments. Since it would be impossible to actually track all of the risk that is created and develop any sort of enforcement mechanism requiring compensation be paid in these cases, a redistributive system based on simplifying assumptions would be more practical and cheaper to operate. Such a redistributive system would notably be somewhat different from the welfare systems that operate in the US or other rich countries, but there would likely be some overlap, at least to the extent that the rich tend to create risk and the poor tend to be put at risk. Fully exploring the claim that poorer people tend to be put at greater risk than richer people is beyond the scope of this paper, but this claim is not obviously wrong. It is, for example, fairly well established that poorer people tend to face a greater risk of death from traffic accidents than do richer people (Harper *et al.* 2015).

I lack the space here to develop such a programme in any detail, and this was not my intention here in any case. I have sought only to show that risk is an overlooked problem for libertarianism and that a plausible way for being able to justify creating risk for others leads us into a libertarian defence of a system of payments based on rectificatory justice.

It might be objected<sup>5</sup> that in many cases we might expect the poor to put the rich at risk. I think, in general, this is unlikely. It should be expected to be the case that as people get more income they choose activities that decrease their risk exposure. So, for example, they buy a car and no longer walk to work. They also might tend to have more political influence, which they use to ensure that a polluting factory or a toxic waste dump does not move into their backyard. Second, it should be expected that as people get more income they will tend to create more risk for others by, for example, driving more, traveling more, and emitting more pollution.

Yet there are at least some cases in which poorer people might create more risk than their affluent counterparts. In some rich countries, many of the poor do own automobiles, and these may be older, and hence emit more pollution, than newer (and more expensive) hybrid or electric cars.<sup>6</sup> It is in such cases that my argument for the libertarian requirement to pay compensation for exposing others to risk is going to diverge from more orthodox defences of redistribution. This is inevitable, since it is unlikely that a form of redistribution based on libertarian principles will entirely align with a form of redistribution based on egalitarianism, or need, or some different justification. Since my argument is about what redistribution libertarianism prescribes, the discrepancy between the libertarian redistribution I have argued here and other forms redistribution does not undermine my case.

A more worrying possibility is whether my argument might be thought to ground a duty of the poor to compensate the rich in cases where they create risk for the rich. There are three reasons to doubt that this actually does pose a problem. First, even if it is sometimes the case that the poor create risk for the rich, for the reasons I have argued above, usually we should expect the relation to be asymmetric in the opposite direction, with the rich creating more risk for the poor. If both groups owe compensatory payments based on risk exposure, the net effect will be that the rich owe more payments to the poor.

Second, the idea of the poor (sometimes) having to pay compensation to the rich is perfectly compatible with our ordinary intuitions about rectificatory justice. If a poor person harms a rich person, it may be perfectly justified to require compensation. To take one example, if a poor person has a bonfire that gets out of control and burns the property of a rich person, it is not clearly false that the poor person owes compensation to the richer because of the harm she caused. Thus, even if the risk principle sometimes requires compensation to go from the poor to the rich, this is not much different from the ordinary intuition that people ought (at least sometimes) to compensate others for wrongfully harming them.

Third, it is possible that compensatory payments might be permissibly postponed or even waived under certain circumstances. Even if we think that the poor person who started the bonfire in the above example ought to compensate their rich neighbour who they exposed to risk (and in this case actually harmed), we are likely to exempt them from paying this compensation if they need that money to survive. While I have not explored destitution as a mitigating factor when it comes to rectificatory justice in this paper, few, including libertarians, would likely object to some form of it. It is hard to defend, for example, requiring a poor subsistence farmer to give all her food and money to another person she harmed or put at risk when doing so would mean that she (and her family) have no food and no means to buy food.

Thus, while the payments based on risk exposure will often flow from rich to poor, there are times where the logic of the argument will require poorer people to compensate richer. I do not think that this is as serious an objection as it might seem since the balance of payments will, for the majority of people, average out in the other direction, in some cases rectificatory justice clearly does require payments from poorer people to richer people, and any reasonable system of rectificatory justice will include some sort of hardship exemption that will ensure that compensatory payments may be put off, reduced, or eliminated when the person required to pay them is facing extreme destitution.

#### IV. CONCLUSION

To summarize, libertarianism faces a large, mostly unacknowledged challenge from risk. It seems to be wrong to foreseeably cause a risk of a harm if the harm itself is prohibited. Rectificatory justice should apply to risk, and not just harm, both because this would prevent too much risk from being produced and because this is required, on deontological grounds, in order to obtain the consent of those placed at risk. In practice, compensation would only be paid if there were institutions in place to monitor risk and enforce payment of compensation. This is identical

in principle to the notion of the night-watchman state, so libertarians who support the night-watchman state should in principle be committed to institutions to monitor risk and enforce compensatory payments. In practice such institutions would be extremely complex, costly, and intrusive, so libertarians ought to prefer institutions instead that use simplifying assumptions and generalizations. This argument grounds a libertarian case for a system of fairly general redistributive payments.

## WORKS CITED

- Hansson, Sven Ove. 2013. *The Ethics of Risk: Ethical Analysis in an Uncertain World*. New York: Palgrave Macmillan.
- Harper, S. T.J. Charters and E.C. Strumpf. 2015. “Trends in Socioeconomic Inequalities in Motor Vehicle Accident Deaths in the United States, 1995-2010.” *American Journal of Epidemiology*. 182/7: 606-614.
- Mill, John Stuart. 1978. *On Liberty*. Indianapolis, IN: Hackett.
- National Highway Traffic Safety Administration. 2018. “Fatality Analysis Reporting System.” <https://www-fars.nhtsa.dot.gov/Main/index.aspx> [accessed July 20, 2018].
- Nozick, Robert. 2013. *Anarchy, State, and Utopia*. New York: Basic.
- Pigou, Arthur. 2018. *The Economics of Welfare*. New York: Palgrave Macmillan.
- Rawls, John. 1999. *A Theory of Justice, rev. ed.*, Harvard: Harvard University Press.

## NOTES

1. I owe this excellent point to an anonymous reviewer.
2. I omit ‘justice-seeking’ here because the libertarian will hold that a justice-seeking social practice will be one that sufficiently respects people’s autonomous choices, and there can, on a libertarian view, be no justice-seeking action that violates other people’s autonomy by exposing them to risks to which they do not consent.
3. I omit ‘*de facto* accept’ here because the libertarian would deny that making use of a benefit means that one *de facto* accepts either paying for it (as in Nozick’s example) or by being exposed to some risk.
4. Pioneered by the economist Arthur Pigou, a Pigovian tax is one that is levied on activities that create externalities for others and is designed to reduce the externality-generating behaviour to a socially optimal level.
5. I owe this objection to an anonymous reviewer.
6. My thanks to an anonymous reviewer for this example.

# *Imposing Risks: A Morally Risky Affair*

Nimrod Z. Kovner

The Hebrew University of Jerusalem, Israel

**ABSTRACT.** In this article, I argue that putting others at risk can establish morally corrupt relations between individuals. The argument follows and advances the writings in legal and moral philosophy that deal with cases of risk-imposition, specifically, the work on the moral significance of risk in normative thinking. However, understanding why imposing risk bears moral significance is crucial for answering related questions, such as what liability or redress means in the context of risk-imposition. Scholars have thus far focused on the impact of risk imposition on individuals' mental state, resources, interests or autonomy. The article's main contribution is its emphasis on the relational aspects of risk-imposition. I claim that putting others at higher risk may form a morally defective relation. I examine two such possible defective relations: (i) domination, based on Pettit's work on Republican freedom and (ii) exploitation, when an agent takes unfair advantage of another agent's vulnerability.

**KEYWORDS.** Risk, domination, exploitation

## I. INTRODUCTION

Consider the case of *Dirty Toys*. Near Little Town, the company Dirty Toys manufactures and sells adorable soft toys but disposes of its toxic waste in a way that contaminates the underground aquifer. The water supply of Little Town is based on the polluted aquifer, from which every resident now drinks. Routine tests of the water quality of the aquifer show high concentrations of toxins, which earlier studies had found to increase the chances of developing cancer by up to thirty percent.

This is a prototype case of industrial pollution, one that unfortunately reminds us of many real-life cases. It is also a classic case of toxic torts.<sup>1</sup> Along with other scenarios where one party exposes others to risk of harm, it raises challenges within legal and moral philosophy. Andrew R.

Klein explains well what makes them hard tort cases: “[l]ong latency periods often precede exposure-related harm, making it difficult for plaintiffs to connect current illness to any specific exposure” (1999, 1174–1175). This Challenge has brought some to entertain the idea of holding agents liable for the risk they impose on others. For the political philosopher Catriona McKinnon, the advantage of liability for risk imposition as opposed to outcomes “[...] is that it can render liable agents who perform acts putting others at risk before the risk materialise, or even if the risk never materialises, and even when those others are unaware of the risks that have been imposed on them” (2012, 78). However, it is left to explain what is morally significant in the risk itself, as a separate matter from the outcomes it denotes, which makes it sensible to hold an agent liable for the mere imposition of risk.

In the present contribution, I follow the writings in legal and moral philosophy that deal with cases of risk imposition. I will not address the question of liability for risk imposition, but will focus rather on the moral significance of risk for normative thinking *per se*. Nonetheless, such an inquiry is important to the issue of liability or responsibility. For the theory of tort law, it will be important to understand what is the moral significance of risk imposition to substantiate and give content to the idea of liability for risk. In moral and political philosophy, such inquiry will help determine why and when imposing risks on others is morally wrong and enrich our thinking about risk regulation. The article’s main contribution is its emphasis on the relational facet of risk imposition, and in forwarding the claim that putting others at higher risk may represent morally defective relations.

I begin with a brief commentary that gives some context to my intervention within the literature on the ethics of risk. From this body of literature, I elicit a core challenge which my paper addresses: namely, to explain moral concerns with risks as an independent matter from the possible future materialised harm that is associated with them. I then present and discuss different strategies to address this challenge. First, I propose the idea that

risk itself can be a source of harm. Risk imposition can have adverse effects, namely psychological harm and the excess cost of protecting against it. I then discuss cases that do not exhibit such types of adverse effects (often called ‘pure risk imposition’). In this section, I present two existing views – the Preference View and the Autonomy View – on the moral significance of risk imposition. In the two sections that follow I advance the main thesis of this essay, demonstrating why, and in what ways, imposing risks is to treat people wrongly. I engage with two concepts – domination and exploitation – that provide content to this more general claim.

## II. SOME PRELIMINARY NOTES

In this section, I offer a brief commentary on the issue of risk ethics. This is not a systematic review.<sup>2</sup> The aim is simply to provide some background to the article, which I hope can help the reader better locate my claims within a certain contemporary discussion.

Let me start by stating one simple way to perceive risk. It is possible to understand *risk* as the “[...] likelihood of a possible negative outcome, such as a loss, injury, harm or death” (Hayenhjelm and Wolff 2012, 29).<sup>3</sup> Risk is sometimes defined more technically in terms of probability of an event, leaving aside the value of such an event. I will focus on the likelihood of bad outcomes and while I sometimes speak about risk in terms of probabilities I still operate within the initial broader notion of risk as likelihood. Though I do not address the distinction between uncertainty and risk, describing risk as the “likelihood of a possible negative outcome” can capture both cases where we can assign probabilities for an event and cases where we do not have knowledge of such probabilities.<sup>4</sup>

Many scholars who write on the ethics of risk share a discontent with the predominant consequentialist framework used to address risk in political theory, public policy, and law. The main target of criticism is Utilitarianism or the method of Risk Cost-Benefit Analysis. The chief charge is that such frameworks lead to conclusions that conflict with the notion

of individual rights and fairness (Hayenhjelm and Wolff 2006, 32-34; Hansson 2003, 293-297). These scholars seek to provide a non-consequentialist alternative that can deal with risk imposition. A central question they deal with is the justification of risk imposition, or in other words, when is risk imposition permissible?<sup>5</sup> Underlying this question is the thought that there is something to be justified; that perhaps there is something wrong or morally bad when we put others at risk. However, here lies the puzzle: what is the moral concern regarding risk itself?

Some of us may share the intuition that since suffering a bad, then causing a bad outcome could be morally wrong. But what is wrong in raising the likelihood of such a bad outcome? It seems that some of the badness of the outcome permeates the risk itself. However, this odd transference of moral value needs to be explained.

To get a better hold of the conundrum, consider the case of *Falling Rock*. I hike in the mountains and reach a cliff, from which I can enjoy a spectacular view. While sitting to rest, I play with some rocks at my feet close to the edge of the cliff. I inadvertently tip a large rock across the edge, and it keeps rolling down toward the path below. A fellow hiker is walking along the path; she is at a high risk of being hit by the rock. As it transpires, the rock falls on the path seconds after she passes, and she continues her walk unharmed.

Since the hiker is not harmed and her trip was not interrupted, where does the problem lie? One way to explain the discomfort that the *Falling Rock* example evokes is to place the problem with the imposition of risk itself.<sup>6</sup> For that, we need to demonstrate the moral significance of risk, or more accurately, of risk imposition. Some scholars have been working on this issue and in the two sections that follow I briefly present their progress.

### III. RISK AND ITS DELETERIOUS EFFECTS

One way to examine the moral significance of imposing risk is to explore its effects. If putting others at risk causes them harm, then such cases

simply fall under our regular moral standards. We are fairly certain regarding the moral significance of harming. There are two ways in which exposing others to risk can harm them: risk exposure can lead to psychological suffering, and it can impose protection costs.

One way in which exposing someone to risk can harm him or her is when the risk becomes an imminent insecurity that hinders his or her ability to enjoy basic freedoms and goods (Hayenhjelm and Wolff 2006). The risk could be taken as a harm, separately from any future potential bad outcome, because of the psychological impact it may have (McCarthy 1997; Meyer *et al.* 2018; Wolff and de-Shalit 2007). Some risks can cause dread and anxiety. The acknowledgment of a looming threat can seep into the daily experience of the individual. Fear begins to accompany and spoil any enjoyable activity he or she participates in. It can be compared to biting into something rotten and not being able to get rid of the taste, which takes over the flavour of any later bite.

Risk can also be harmful in an additional way: knowing that you are subject to the chance of suffering a bad outcome compels you to invest in protective measures that will avert the harm or soften the blow if it comes to pass (Wolff and de-Shalit 2007). Therefore, we can say that putting you at risk can lead to additional costs and a diversion of resources from objectives and plans you have. In this form, putting others at risk is to harm them because of the additional costs imposed on them.

Both harmful risks – psychological suffering and excess protection costs – have a significant shortcoming. Without any knowledge of the looming threat, one cannot suffer psychological harm or try to avert the risk (and incur the associated costs). However, in many cases where people are oblivious to the potential danger, it is still reasonable to hold that the exposure to risk makes their situation worse than it otherwise would have been. Such scenarios of risk imposition raise the need to explore other ways, beyond such clear cases of harming, in which imposing risk still bears a moral meaning that we should address.

## IV. PURE RISK IMPOSITION

A useful way to explore the moral significance of risk imposition beyond causing tangible harm is to analyse cases in which such harms do not feature at all. Some use the term ‘pure risk imposition’ to describe such cases.<sup>7</sup> Suppose that one agent creates a risk to another agent without her awareness and without any indication that the risk will either pass or materialise.<sup>8</sup> To make the case neater, some assume that the risk has passed without materialising into an actual harm. In such cases, the primary challenge lies with those wishing to protect against risks or compensate for it.

To better understand the task ahead, let us take a new example *Reckless Driving*.<sup>9</sup> Ellie finished her work for the day and is heading home. She gets into her car, starts the engine and speeds away. She is late for dinner with her family so she is in a hurry and she decides to drive faster, far beyond the official speed limits, passing other cars and dangerously crossing lanes. Fiona is also driving back from work. She shares part of Ellie’s route and will pass it at the same time. Due to Ellie’s reckless driving, Fiona is much more likely to be involved in a car accident that day.

Ellie’s reckless driving exposes Fiona to risk that is higher than the risk of her usual commute, even if she is not aware of it and will not end up crashing her car. In other words, Fiona does not suffer any tangible harm. Nonetheless, I think many of us will judge Ellie’s driving as morally wrong. But what exactly is wrong with her action? We find two main possibilities in the literature, which can be called the Preference View and the Autonomy View.

According to the Preference View,<sup>10</sup> best represented by Claire Finkelstein (2003), cases of pure risk imposition can also be treated as harming others. She bases her view on Joel Fienberg’s influential conceptualisation of harming others that understands harm as a setback to legitimate interests (Finkelstein 2013, 971). Since this interpretation of harm is fairly rooted in our moral understanding, showing that pure risk imposition falls under this

discription provides us with its moral significance. According to Finkelstein, in a similar fashion to the way agents *prefer* to be in a group that is offered a higher likelihood of gains (be part of a lottery), they prefer not to be part of a set of people with a higher likelihood of harm (risk).<sup>11</sup>

In *Reckless Driving* it is clear that Fiona has an interest in a safe journey home; she wants to complete her drive without being harmed. If she were to be asked, she would state her preference not to be one of the drivers on Ellie's path. While acknowledging that driving involves risks of harm, Fiona would not prefer to be exposed to the higher risk created by Ellie's reckless driving. For the Preference View, Ellie harms Fiona by imposing on her such excess risk.

The Autonomy View is associated with the work of John Oberdiek. Oberdiek (2012; 2017) argues that by putting other agents at risk, at present, we are changing their map of possible futures in terms of their set of possible options, their likelihood, and their quality:

Imposing risk does just this in virtue of narrowing the risked person's otherwise 'open future'. [...] One can do harm in subjecting a person to risk, for it effectively attaches sanctions to or normatively forecloses certain options that would otherwise be available to the individual, thereby narrowing the risked person's set of worthwhile opportunities (2012, 351-352).

Oberdiek (2017) explains the moral significance of risk in such cases in terms of curtailing the agent's autonomy, using Joseph Raz's influential account of autonomy. When we assess an agent's state of affairs, the valuable options she has are an important measure. If someone blocks one of her valuable options, our assessment will change. We would judge the new state of affair as worse than the previous one. In some sense, the risk of harm squanders the value of an option because part of what makes it worthwhile is the ability to enjoy it safely.

In the case of *Reckless Driving*, we can say that Fiona's route home, although it was never free from risk, has become more dangerous as a

result of Ellie's actions. Ellie removes Fiona's option of a relatively safe drive, which she had before Ellie decided to speed home. Now, there are fewer possible futures in which Fiona returns home safely.

One might insist, however, that it is still unclear, in *Reckless driving* for example, what is the moral significance of Ellie's actions on Fiona's overall state of affairs. Ellie alters Fiona's situation, that seems true, but in what has been said thus far there is not enough to show in what way her situation is worse than before the intervention. To write that she becomes 'worse off' due to the risks imposed by Ellie's driving is insufficient as an explanation. The way in which she is worse off is what needs to be explicated.

So, let me recap on what we know and what is left to answer before moving forward with the argument. Fiona has fewer safe routes on her commute; she cannot obtain the same risk level of driving back home as before. However, if she is not physically or psychologically harmed, if she does not change her plans, what gives substance to the intuition that her situation has worsened? According to the Preference View, Fiona would not prefer to run this risk. However, it seems that the content of this preference is lacking. When we postulate that Ellie's reckless driving has no tangible possible effects, it is not clear what the reasons are not to prefer it over a scenario in which Ellie drives responsibly.

Oberdiek thinks that such risk imposition curtails an agent's autonomy, which is morally significant but not wrongful from the outset. According to Oberdiek it is a "candidate wrong" (2017, 91), which means that the wrongness of such risk imposition will be determined according to how and if we can justify it. The Autonomy View only explains what the morally relevant feature of risk imposition is; the complete moral judgment rests on the justification (or its lack thereof) of risks. While I accept this strategy, I also think that neither the Autonomy View nor the Preference View provide us with all the morally relevant features of risk imposition. The removal of valuable possible options or the justified preference against risk do not exhaust the moral concerns regarding risk imposition.<sup>12</sup>

We should also pay attention to the relational dimension of pure risk imposition. In other words, I argue that by imposing a risk on others we may treat them wrongfully, even without causing them material harm.

## V. IMPOSING RISK AND NON-DOMINATION

The first type of potentially wrong kind of relations that risk imposition expresses is *domination*. Interestingly, Oberdiek himself comes close to the language of domination at some point when he elucidates his autonomy-based argument. He writes that “[...] imposing risk on others is morally significant in these terms because risk impositions amount to claims of authority over others’ lives” (2017, 90). In a broad sense, domination is to have the power to decide on others. It should not be surprising that autonomy and domination are understood in similar terms, as both are predominantly about freedom. However, the concept of domination adds a relational dimension to our understanding of freedom, one that is missing from the concept of autonomy. The notion of autonomy Oberdiek alludes to is based on the quantity and quality of available options; what the agent ‘possesses’. Domination brings in power relations; the power to decide for another which options will be available and which options will not.

Domination (or more accurately its absence – non-domination) is central to the Republican notion of freedom. Republican freedom stresses that actual interference with an agent’s choices, which can be understood as coercion, cannot account for the whole meaning of what it is to be free.<sup>13</sup> We must also consider *possible* constraints on options to determine if one is free or not. When one agent can potentially constrain another agent, she obtains power over him, even if she does not use it. The agent subjected to such decisional power is being dominated, even when there are no actual limitations to his choices at a given moment. Freedom, under this view, is about non-domination, changing the power relation that subjects agents to the unchecked authority of others.

Philip Pettit famously developed this Republican theory of freedom. According to Pettit, A is unfree when B has the capacity to interfere, on an arbitrary basis, with choices that A is in a position to make (1997, 52). Two main components of the theory distinguish it from the standard view of freedom and its focus on actual impediments: the possibility of interference and its arbitrariness. What counts as arbitrary interference can be debated.<sup>14</sup> For Pettit, “When we say that an act of interference is perpetrated on an arbitrary basis, then, we imply [...] that it is chosen or rejected without reference to the interests, or the opinions, of those affected” (1997, 55). The possibility of interference is an easier point to grasp. We should also account for possible hypotheticals in which such intervention occurs. They show us that while it now seems that an agent is free to choose from his available options as he pleases, he is subject to the decisional authority (or ‘mercy’, to use Pettit’s dramatic tone) of another agent.

Now let us go back to risk imposition and examine how the concept of domination may help us understand their moral significance. In *Reckless Driving*, Ellie imposes on Fiona a greater risk of injury. We can imagine possible scenarios in which Fiona (or other drivers on Ellie’s route) is involved in a car accident. For example, Ellie’s briefly loses concentration when crossing lanes and she does not see Fiona’s car coming towards her; as a result, the two collide. An accident is not a remote possibility considering Ellie’s driving and Fiona’s timing. This possibility is by itself an interference with the valuable choices Fiona can make, namely with getting home safely. Ellie’s driving represents an arbitrary intervention, as it does not track Fiona’s interests, fails to express adequate consideration for her life, while having the potential to harm her. In a way, by generating risks for Fiona (and other drivers), Ellie places a wager on Fiona’s future. She takes some significant aspect of Fiona’s life into her hands, so to speak. Such power over others’ lives, to my mind, is domination.

But not all acts of creating risk to others create relations of domination. Here we can again draw on Pettit’s work. He writes that “[...] an

act of interference will be non-arbitrary to the extent that it is forced to track the interests and ideas of the person suffering the interference” (1997, 55). If the activity generating the risk must give proper weight to the interests of those it subjects to the risk, relations will not be of the dominating kind. It is hard to see this in the case of *Reckless Driving* because we have an existing standard for safe driving that clearly makes Ellie’s conduct wrong in the sense described here. In this scenario, it is hard to think of a way to track Fiona’s interests that would not simply demand that Ellie drive with greater care; that is, to avoid imposing the enhanced risk all together.

So, think of a different case of fast driving: *Undercover Operation*.<sup>15</sup> Jane Smith is a secret service agent working undercover. She receives an intel regarding a time bomb that is set to explode in a shopping centre on the other side of town. She cannot call for backup as the terrorist will override the timer and trigger the bomb manually if he sees the police arrive. Jane knows she can prevent the explosion if she gets to the shopping centre on time. She runs to her car and drives as fast as she can; she exceeds above the speed limit, crosses lanes, passes other cars and even ignores the lights. Thankfully, she does not hit any other car and arrives just in time at the shopping centre to disarm the bomb before it goes off.

Our thoughts on *Undercover Operation* and *Reckless Driving* will inevitably differ. While we cannot see how Ellie’s driving can trace or give adequate weight to Fiona’s (or any other driver on the road) interests, Jane’s drive may fare better by this measure. This specific drive to the shopping centre may not be in the current immediate interest of the drivers it puts under the higher risk of an accident. However, as a rule, we would accept that during rescue missions, like Jane’s operation, agents will ignore the regular traffic regulations. It is in everyone’s interest, as everyone would want a member of the security forces to reach a bomb located next to them and disarm it as quickly as possible.

For Pettit, part of the problem of arbitrary interference is the unchecked power one has over the other. Therefore, in Pettit’s theory,

we should strive to balance the ability of individuals to exercise power over others in order to achieve an ideal society of equals. But for my analysis here, what is important is how the concept of domination exposes a relational dimension of risk imposition. The morally relevant features of *Reckless Driving* go beyond the preference of Fiona not to be exposed to higher risk of an accident, or the quality of the available option-set she can obtain. Something important about how Fiona's life will go forward is decided by Ellie, without any regard to Fiona's interests and without her say on the matter or ability to influence the decision. In short, imposing risk forms power relations. Not all such relations will be judged as morally wrong (as I suggested with *Undercover Operation*), but it shows that the way individuals are treated by being put at risk can be morally significant.

## V. IMPOSING RISK AND EXPLOITATION

Exploitation can have different meanings and philosophers use it in varied ways in normative arguments. In the broadest sense, “[t]o exploit someone is to take unfair advantage of them. It is to use another person's vulnerability for one's own benefit” (Zwolinski and Wertheimer 2017). I will not attempt to provide a well-defined concept of exploitation that can be applied to all cases involving risk. What I will do, is explicate some features of exploitative relations pertinent to risk imposition that show its moral significance.

Placing a bet on the life of another individual could be a morally questionable way to treat that person in a different way than domination represents; it can be exploitation. Exploitation is for most theorists a moralised concept, in the sense that once relations are identified as exploitative we regard them as wrongful, at least if there are no countervailing reasons to judge differently (Zwolinski and Wertheimer 2017). In that sense, exploitation differs from domination. However, the two concepts are not far apart. In fact, we can think of exploitation as a subset of domination,

where all exploitative relations are a form of domination but not all domination cases are exploitative relations.<sup>16</sup> We can observe this if we look at the structure of exploitative risk imposition. On one side we have a risk-taker who is also the beneficiary of the risky activity. On the other side we have those under the generated risk, the risk-imposed.

Recall the example of *Dirty Toys*, where the residents of Little Town were exposed to toxins that increase the probability of developing cancer. The company polluted the source of drinking water because it is a cheap waste disposal method. The company decided to engage in an activity that saved money while exposing others to serious health risks. Here, the company is the risk-taker and the beneficiary and Little Town's residents are the risk-imposed. It is an intuitive case of exploitative relations, where the company takes advantage of the relative vulnerability of the residents to benefit itself. The distribution of risks and benefits is neat and clear.

There are, of course, messier cases of risk imposition. We may have cases where the risk-imposed may benefit from the risky activity, but the risk-taker enjoys both safety and a greater share of the benefits. This may still be a case of exploitation, as it is possible that both parties will be better off than they would have been without the risky conduct, yet one party will still gain unfair benefits from it. Much of the debate over the concept of exploitation revolves around such relations. It is not easy to decide if there is exploitation or some form of wrongful treatment in many cases of a mutually beneficial transaction with an unequal share of benefits. I think that much depends on the distribution of risks and benefits and the power relations operating in the background.<sup>17</sup> But for the purposes of this article, what is important to emphasise is that in order to form exploitative risk imposition, the risk-taker must also be the beneficiary. This is not required for domination cases (those that are not also exploitative relations). There could be domination cases in which there are no clear benefits (or comparatively greater benefits) for the risk-taker.

The dual role of risk-taker and beneficiary also distinguishes exploitative risk imposition from other kinds of exploitative relations mentioned

in the literature. There are types of exploitation in which the beneficiary is the exploiter, but she is not the reason other individuals are vulnerable. Benefiting from past injustice (assuming no causal responsibility for such injustice) and free-riding could be examples of this kind of exploitation. Such cases probably have morally relevant features; they might be exploitative or morally defective relations for other reasons. However, my focus is on the moral significance of risk imposition, a situation in which an agent generates risk for others. Therefore, cases in which an agent is merely the beneficiary of others' vulnerability are beyond the scope of this investigation.

There are other distinctions regarding exploitation that can assist in clarifying the more specific type of exploitation that risk imposition expresses. There is *substantive* exploitation, for example, where the outcome or risk imposition is in some sense unfair. This is opposed to *procedural* exploitation in which there is something not right in the process that leads to unfair outcomes (Zwolinski and Wertheimer 2017). Of course, with risk imposition, talking about 'outcomes' is more complex, and I address it below. Risk imposition falls under *transactional* exploitation as opposed to *structural* exploitation. The latter is about social systems or social institutions that unfairly benefit some at the expense of others. Transactional exploitation is about relations between two (or more) agents, where one agent (or more) takes unfair advantage of the other agent (or more) (Zwolinski and Wertheimer 2017). Most examples of transactional exploitation present agents interacting with each other: selling something, making an offer, a contract, or within labour and production relations. Exploitative risk imposition could manifest differently without such direct or ongoing interactions. Also, the benefits of exploitative risk imposition do not have to flow from the risk-imposed to the risk-taker. Benefits can be extracted from the risky activity and not from the activity of the exploitee, as with someone who is able to pay another a meagre sum for her to do a valuable task. Think once more of the case of *Dirty Toys*. The company creates an environmental hazard that

jeopardizes residents' health without direct interaction with them. The company benefits because it saves money by using a dangerous but cheap waste disposal method. Dirty Toys company does not benefit from the actions or choices of Little Town residents whom it puts at risk.

Following this brief discussion, we can spell out two necessary conditions for exploitative risk imposition: (i) agent (or agents) creates a risk for other agent, and (ii) she is the beneficiary of the risk-generating activity or unfairly benefits more from it. Direct interaction between the agents is possible but not necessary, and benefits can, but do not have to, flow from the risk-imposed to the risk-taker.

Let me spell out what this notion of exploitation adds to existing accounts of the moral significance of risk imposition? In *Dirty Toys*, the residents of Little Town are unaware of the higher risk they are exposed to because of the contaminated water resource. Therefore, they do not suffer the emotional strain of living under risk. They also do not invest in precautions against the risk; they do not undergo costly medical examinations, for example, or buy bottled water. In other words, they do not incur the excess cost of risk avoidance. Described this way, *Dirty Toys* is a case of pure risk imposition. The Preference View would claim that the residents would not choose to be exposed to such risk; it is against their interest in good health. The Autonomy View will claim that the residents' option to use safe drinking water was removed by the company; the autonomy of every individual was thus compromised. In order to isolate the relational aspect of risk imposition, it will be useful to assume that the risky situation has passed without materialising into harm.

So, consider the possibility that the *Dirty Toys* company stopped polluting, the toxic water reservoir was pruned, and no one in Little Town developed cancer. Now, when all are safe and well, we ask again what is our moral concern with the story of *Dirty Toys*? Since the life plans of all individuals continued without any impact from the pollution, it is not clear how their autonomy was compromised. In retrospect, we see that developing cancer was a remote possibility with minimal effect, if any,

over residents' options. Also, it is not clear what the content of a preference not to be exposed to risk when no harm comes about in the end. When we clear away the harmful outcome from our examination of the case, we have less confidence in the Autonomy and Preference Views' argument for what is the moral significance of pure risk imposition. However, the way in which the company treated the residents continues to resonate in their relationship even after the risk has faded without any harmful effect. While no one is harmed, the company still advanced its economic interest at the expense of the quality of Little Town's drinking water. It took advantage of the residents' lack of awareness to its disposal method and their resulting incapacity to prevent it (or oppose it). In short, the company exploited them. Exploitation, then, is one way to describe the relational moral feature of such risk imposition. Moreover, such exploitative risk imposition is, at the very least, *pro tanto* wrong.

## VI. CONCLUSION

Risk is ubiquitous. To impose and be exposed to risk is part of how we live with others. This gives us enough reason to think about the normative meaning of risk both in regard to our own actions as well as the rules and regulation governing society. In the present contribution I presented some of the progress made in this philosophical endeavour. As opposed to actual cases of harming others, cases of risk imposition face theoretical challenges. One central challenge is to understand what is morally wrong with the risk itself. After all, a risk is merely the likelihood of a bad outcome.

I presented several strategies to overcome this challenge. I first showed that risk imposition can cause psychological suffering; living with risk can be mentally strenuous, generating feelings of fear, dread, and anxiety. Also, risk exposure can impose additional costs of protecting against it; investing resources in averting the danger, fencing against it, or developing the capacity to cope better with the bad outcome. I then discussed cases of

‘pure risk impositions’, in which risk-as-harming is out of the picture. I presented the Preference View claiming that agents have a justified preference not to be put at higher risk, and the Autonomy View that depicts imposing risk as the elimination of valuable possible options. Without rejecting these two views, I stressed that they lack an important aspect of risk imposition. Both the Preference and the Autonomy Views focus on changes to the status of individuals at risk and they do so in isolation from the fact that this risk is imposed on individuals by others. They use the word ‘imposed’, but they do not address the moral significance of risk imposition as an action taken by an agent. In other words, I suggest examining how imposing risk may be a wrongful way to treat people.

The relational aspect of risk imposition is not necessarily a rival to other morally relevant features reviewed in this essay. But in order to understand this aspect it is useful to examine it separately. Isolating the relational feature from other morally relevant features (such as suffering psychological harm or excess protection costs) serves two functions. First, it helps see other morally relevant features of risk imposition more clearly. Second, it helps us to understand the moral significance of risk imposition in cases where some morally relevant features are absent (when those at risk are unaware of the risk, for example). With this strategy in mind, the relational account I suggest has some advantage over the Preference and Autonomy Views, because the way individuals are treated is a persistent feature of risk imposition. When risks do not ripen into harm, the impact of the potential harmful outcome on our moral judgment dissolves. But it also weakens the power of the non-tangible effects that the Preference and Autonomy Views propose. However, we can be more certain about the relational state of affairs – putting others at risk remains a feature of and a possible taint on the relationship.

I used two concepts – domination and exploitation – to examine possible types of risk relations, both of which give rise to moral concerns. Domination is a morally relevant feature of risk imposition as it implies the exercise of decisional authority over others. I suggested that Oberdiek’s

Autonomy View is not that far from the notion of domination; both address the issue of freedom. However, Oberdiek's account remains focused on the set of possible valuable options an agent faces. In this sense, his account is too 'arithmetic' and neglects the moral significance of being under the authority of another; how one can dominate another by risk imposition. I think that being under the power of others presents a clearer and stronger offense to moral values than the way Oberdiek suggests risk imposition compromises autonomy. In this sense, it can provide a fuller and more robust explanation to Oberdiek's Autonomy View.

I also suggested that putting other at risk in order to extract unfair benefits forms exploitative risk imposition relations. Exploitation, once identified, can be judged as wrong, at least *pro tanto*. Exploitation provides a good account of the moral significance of pure risk imposition. Even without harming those exposed to risk, and sometimes even when they are in some sense better off, the risk-taker can unfairly benefit at the expense of those she puts at higher risk. I think the notion of exploitation is one good way to put flesh on the bones of the Preference View's more general claim. In cases of pure risk imposition, where we assume that the harm will not, or even has not materialised, it may not always be clear what is the content of the preference not to be exposed to risk. However, it is reasonable to have a strong preference not to be exploited by others, even under such an assumption. This is simply not how we want to be treated, and more importantly not how we should be treated.

With this exploration of the moral significance of risk relations, I add to the developing field of risk ethics. Uncovering the moral meaning of putting others at risk is important to other questions regarding the ethics of risk. What risks are permissible? How can we justify risk imposition? And what should redress or compensation for risks look like? To provide good answers to such questions, we need to know what is morally wrong with risk imposition or at least what is morally significant in such cases. For example, if domination is a morally relevant feature of risk imposition,

then risk justification must take it into account. Since the core concern of domination is the arbitrary use of power over others, then the justification of risks should also consider avenues of influence over the relevant decisions, or form regulation guarding against the unchecked power to impose a risk on others. Likewise, to justify situations suspected as exploitative risk imposition, it is necessary to address the distribution of risks and benefits. These examples indicate how, through advancing our understanding of the moral meaning of risk imposition, my present contribution assists in the proper formulation and analysis of the ongoing philosophical inquiry into the ethics of risk.<sup>18</sup>

#### WORKS CITED

- Arneson, Richard. 2013. "Exploitation and Outcome." *Politics, Philosophy & Economics* 12/4: 392-412.
- Aven, Terje. 2010. "On How to Define, Understand and Describe Risk." *Reliability Engineering and System Safety* 95: 623-663.
- Carnor, C. F. 2007. "Toward a Non-Consequentialist Approach to Acceptable Risks." In *Risk: Philosophical Perspectives*. Edited by Tim Lewens, 36-53. London: Routledge.
- Ferguson, Benjamin. 2016. "The Paradox of Exploitation." *Erkenntnis* 81: 951-972.
- Finkelstein, Claire. 2003. "Is Risk a Harm?" *University of Pennsylvania Law Review* 151/3: 963-1001.
- Hansson, Sven Ove. 2003. "Ethical Criteria of Risk Acceptance." *Erkenntnis* 59: 291-309.
- Hansson, Sven Ove. 2004. "Philosophical Perspectives on Risk." *Techné: Research in Philosophy and Technology* 8/1: 10-35.
- Hayenhjelm, Madeleine and Jonathan Wolff. 2012. "The Moral Problem of Risk Impositions: A Survey of the Literature." *European Journal of Philosophy* 20/1: 26-51.
- Klein, Andrew R. 1999. "A Model for Enhanced Risk Recovery in Tort." *Washington & Lee Law Review* 56/4: 1173-1210.
- Kumar, Rahul. 2015. "Risking and Wronging." *Philosophy and Public Affairs* 43/1: 27-51.
- Lenman, James. 2008. "Contractualism and Risk Imposition." *Politics, Philosophy, and Economics* 7: 99-122.
- List, Christian and Laura Valentini. 2016. "Freedom as Independence." *Ethics* 126/4: 1043-1074.
- McCarthy, David. 1997. "Rights, Explanation, and Risks." *Ethics* 107: 205-225.
- McKinnon, Catriona. 2012. *Climate Change and Future Justice: Precaution, Compensation, and Triage*. London: Routledge.

- Meyer, Lukas H., Fabian Schuppert, Harald Stelzer and Adriana Placani. 2018. "Risk and Rights: How to Deal with Risks from a Rights-Based Perspective." In *Towards the Ethics of a Green Future: The Theory and Practice of Human Rights for Future People*. Edited by Marcus Düwell, Gerhard Bos and Naomi van Steenberg, 28-47. London: Routledge.
- Oberdiek, John. 2012. "The Moral Significance of Risking." *Legal Theory* 18/3: 339-356.
- Oberdiek, John. 2017. *Imposing Risk: A Normative Framework*. Oxford: Oxford University Press.
- Perry, Stephen. 2007. "Risk, Harm, Interests, and Rights." In *Risk: Philosophical Perspectives*. Edited by Tim Lewens, 190-209. London: Routledge.
- Philip, Pettit. 1997. *Republicanism: A Theory of Freedom and Government*. Oxford: Clarendon.
- Schroeder, Christopher H. 1990. "Corrective Justice and Liability for Increasing Risks." *UCLA Law Review* 37/3: 439-478.
- Teuber, Andreas. 1990. "Justifying Risk." *Daedalus* 119/4: 235-254.
- Thomson, Judith J. 1986. *Rights, Restitution, and Risk*. Cambridge, MA: Harvard University Press.
- Zwolinski, Matt and Alan Wertheimer. 2017. "Exploitation." *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. <https://plato.stanford.edu/archives/sum2017/entries/exploitation/> [accessed July 11, 2018].
- Wolff, Jonathan and Avner de-Shalit. 2007. *Disadvantage*. Oxford: Oxford University Press.

## NOTES

1. The hypothetical example is constructed along the lines of two well-known American court cases: *Ayers v. Township of Jackson*, 525 A.2d 287 (NJ 1987); *Sterling v. Velsicol Chemical Co.*, 855 F.2d 118 (6<sup>th</sup> Cir. 1988).
2. For those interested in such a review, Hayenhjelm and Wolff (2012) makes a good read.
3. Risk can be understood in other ways as well. For a discussion of different definitions of risk, see Hansson (2004); Aven (2010).
4. It is customary in some fields of inquiry to follow the distinction introduced by Frank Night, which reserves the term 'risk' for cases in which we can ascribe probabilities, and the term 'uncertainty' for cases with unknown probabilities. In our everyday language we quite often conflate the two and generally take uncertainty to signify some epistemological incompleteness.
5. On justifying risk imposition, see Carnor (2007); Kumar (2015); Lenman (2008); Oberdiek (2017); Teuber (1990).
6. It is possible that our discomfort in this case is too heavily based on the severity of the outcome – getting injured by a falling rock. It is hard to illustrate cases of risk imposition without invoking some of the reader's biases. And it is hard to isolate our thoughts regarding the risk itself from the outcome it denotes. In *Falling Rock* the fact that the rock misses the hiker aims to help us overcome our typical fixation with outcomes. We are left freer to contemplate on what is

morally troubling in placing the hiker at such risk of injury (I thank an anonymise reviewer for pressing me to make the purpose of this example clearer).

7. Judith J. Thomson (1986) has coined the term ‘pure risk imposition’, which others use as well.

8. Stephen Perry (2007) objects to the notion of ‘pure risk imposition’ as harming others and claims it may still be wrongful on other grounds. My argument stands regardless of the position one holds in this debate, as I bring forth another moral dimension of risk imposition; arguing for the moral significance of the relations risk imposition forms.

9. This example is used, with some variation, by different philosophers. For one example, see Schroeder (1990).

10. The title follows Oberdiek (2017) and Meyer *et al.* (2018) who use Preference-Based Account to refer to Finkelstein’s position.

11. For criticism of Finkelstein’s Preference View, see Perry (2007); Oberdiek (2017, Chapter 3).

12. The Autonomy and Preference View are not necessarily compatible accounts as I present here. Oberdiek presents them as competitive views, for example. I will not enter this debate here. My argument stands even if these views, together or apart, are true or false.

13. It is important to note that for most political philosophers, freedom is understood normatively. By that I mean that the freedom they are interested in is the absence of agential constraints (actual or possible) and not other non-agential inhibition (like a natural disruption) to our choices.

14. For example, see List and Valentini (2016) criticizing Pettit’s view of arbitrary interference as a ‘moralised’ conception of interference.

15. I thank an anonymous reviewer for suggesting this example.

16. For an elaborated argument on the relation between economic exploitation and domination, see Vrousalis (2013).

17. For discussions on the challenge of mutually beneficial transactions and exploitations, see Arneson (2013); Ferguson (2016); Zwolinski and Wertheimer (2017).

18. A previous draft of this article was presented at the doctoral and postdoctoral workshop at the Hebrew University Law Faculty. I wish to thank the participants, and in particular Merav Kaddar, Re’em Segev and Ori Herstein for their helpful comments on the workshop article. I would also like to extend a special thanks to Kai Spiekermann for his ongoing guidance and support and to Nourit Zimerman for proofreading the article.

# *Justice and the Distribution of Risks from Unconventional Energy Technologies*

Aviezer Tucker  
Harvard University, US

**ABSTRACT.** The application of the theories about risk distribution of Beck, risk assessments of Luhmann and Giddens, and justice as the internalization of externalities of Schmitz to the analysis of the social and political reactions to new ‘unconventional’ technologies for extracting natural gas and oil from shale rock discovers correlations between injustice in the distribution of risks and rewards and risk assessment and trust, reinforced by a feedback causal loop in which high assessment of risk and distrust reinforce each other. The unaccountability of technocratic and political elites has increased distrust and generated a feedback causal loop that increases the unaccountability of elite decision making when elites adjusted to popular distrust by attempting to enact decisions about the distribution of risk from new technologies in stealth or when the population was distracted. Geopolitical risk may balance technological risk assessment when the assumption that technological risks reduce geopolitical risks. The article applies this model to explain the politics and policies of unconventional energy in Europe. It further supports the model by a comparison with regulatory policies for the extraction of unconventional energy resources in the United States and by examining a case in England, where the government attempted to decrease injustice in the distribution of risk.

**KEYWORDS.** Beck, Ulrich, justice, energy, hydraulic fracturing, risk, trust

## I. INTRODUCTION

The present contribution argues that justice in the distribution of risks, risk assessments, and trust in expertise and technocratic and political elites are closely connected. Distributive justice is not only a normative theory that should regulate the distribution of risks and rewards, it is also a political force that affects policies. Those who benefit from the risks others are forced to assume tend to underestimate risks; and vice-versa, those who are unjustly deprived of the rewards of the risks they are

forced to assume overestimate the degree of risk to which they are exposed. At the same time, the victims of injustice in the distribution of risk come to distrust those who benefit from the unjust distribution and those who aid them, or seem to support them, by underestimating or at least not overestimating the risk, namely technocratic and political elites and experts. The effects of the injustice in the distribution of risk, overestimation of risk and distrust of elites and experts, reinforce each other. Victims of injustice who distrust experts have no alternative authoritative source for their own higher risk estimates. Vice versa, overestimation of risk leads to distrust of technocratic elites and experts who disagree with that estimate, especially if it is expressed condescendingly. Schematically, the theory is represented in the following flow chart:



I test this theory against, and use it to explain, political processes and policies about new extraction technologies of unconventional energy in Europe.

First, I outline the theory following its origins in philosophy and social theory in David Schmidtz's (2006) theory of justice and the social theories of risk proposed by Beck (1992), Luhmann (1993) and Giddens (1990). I then explain the risks that accompany new technologies for extracting unconventional energy resources and outline policy responses to them in Europe. I apply the theory to explain the political reactions and policy results in Europe. Since the social world is not a laboratory in which the three variables interact in isolation, a number of other variables have affected political debate and resulting policies concerning unconventional energy extraction in different countries. I highlight a feedback causal loop between unaccountable technocratic and political elites and

distrust. Geopolitical risk may mitigate the perception of risk from new technologies. Different distributions of risks and rewards in the US further support the theory. Finally, I explore experiments with reducing injustice in the distribution of risks and rewards from the exploitation of unconventional energy in England.

## II. THE THEORY OF UNJUST RISK DISTRIBUTION, RISK ASSESSMENT, AND DISTRUST

Justice regulates the distribution of risks no less than the distribution of goods and opportunities. David Schmidtz (2006) proposes a particularly appropriate theory of justice for regulating the distribution of risk. Schmidtz interprets justice as the *internalization of externalities*. Actions have consequences that should be paid for by the agents who chose to undertake the action and not by others who are affected without choice. For example, if some people choose to burn fossil fuels that generate pollution and they profit from it, it is unjust that other people should suffer the consequent pollution; the external costs should be internalized. Applied to risk, internalizing externalities means that risk managers who decide to support risky policies should have to bear the risk rather than transfer it to others, let alone divide risks from rewards and externalize the former by transferring it to others while appropriating the possible rewards. If some people are forced to bear risk they should be compensated for it and receive the rewards if their assumption of risks generated them. Unjust distribution of risk has two mutually reinforcing effects: it heightens the perceptions of risk among those forced to bear it without receiving potential rewards, and it destroys their trust in those who externalize the risks, namely elites and experts.

Luhmann (1993) distinguishes risk, or the future adverse result of *decisions*, from *dangers*, or adverse effects that do not result from decisions. For example, there is a *risk* of global warming, but a *danger* that changes in the sun may cause changes in earth's climate. Beck (1992)

argues that the distribution of risk in society has become politically more contentious and significant than the distribution of social goods like wealth. Risks that are distributed unequally and without corresponding rewards become political issues. Since risk results from decisions, those who make the decisions may be distinct from those affected by them, i.e. those who pay the cost of externalities. This social division has generated social reactions and protest movements among those forced to assume externalized risk by the decisions of others. The division between those who take risky decisions and those affected by them is based on, and also leads to, distrust of the former by the latter. Though decision makers attempt to ‘reason’ with those who are asked to accept the risk, they are rarely convincing because they are not trusted. “[T]he conflict between decision makers and affected parties cannot be influenced by quantitative analysis of risky situations” (Luhmann 1993, 148). Unjust distribution of risk in society begets a heightened perception of risk by those forced to bear it:

What is to count as a disaster is not decided on the basis of objective criteria [...] Conceding the circularity of the argument, we shall therefore speak of disaster whenever the affected party refuses to allow himself to be convinced by quantitative analysis. This disaster threshold is set in very different ways by the politically relevant population and above all by the mass media [...] This means that politics cannot rely on the quantitative calculation of risk and that it cannot be expected to do so. Instead it has to make do with informal guesses on the effects and above all on the acceptability *of its own decisions* (Luhmann 1993, 149-150; italics original).

Giddens characterizes modernity as founded on trust in expert systems. “Trust may be defined as confidence in the reliability of a person or system, regarding a given set of outcomes or events, where that confidence expresses a faith in the probity or love of another, or in the correctness of abstract principles (technical knowledge)” (1990, 34). Luhmann (1993) and Giddens (1990) connected the *perception of*

*risk* with the *absence of trust*. People who are forced to involuntarily bear risk without a corresponding greater share of the potential rewards distrust those who force this unjust distribution on them. Lay persons develop a cynical approach to experts whom they do not trust but cannot avoid.

Experts often take risks ‘on behalf’ of lay people while concealing, or fudging over, the true nature of those risks or even the fact that there are risks at all. More damaging than the lay discovery of this kind of concealment is the circumstance where the full extent of a particular set of dangers and the risks associated with them is not realized by the experts. For in this case what is in question is not only the limits of, or the gaps in, expert knowledge, but an inadequacy which compromises the very idea of expertise (Giddens 1990, 130-131).

Experts conceal much of their ‘backstage’ activities from the public and often prefer to reassure lay people by concealing how often contingency and luck are involved in expert actions in order to maintain trust (Giddens 1990, 86). When experts are socially and politically diverse, forming a heterogeneous community, they may claim that their professional consensus about risk assessment reflects their shared knowledge (Tucker 2014a). However, when they appear like a homogeneous social class, their agreement on low risk assessment may serve as evidence of their biased political or economic interests. The gap in risk assessment between experts and victims of injustice in the distribution of risk reinforces distrust in experts as intellectual mercenaries, which in turn reinforces the gap in risk assessments; the two variables causally loop on each other, as they follow independently unjust risk distribution.

States and technocratic elites may attempt to diffuse social tensions between those who decide on and benefit from risk distribution, and those who pay its external cost through ‘absorption strategies’ such as participative democracy, organized/managed communication, and ethical technology assessment. Such strategies force the opposition to

participate in decision making and to justify itself (Luhmann 1993, 152). Communication between risk managers and risk bearing locals, however, cannot change their different risk assessments. At most, it can correct errors. Distrust remains along with suspicion of the manipulative organizers of deliberative forums, such as deliberative polling, consensus conferences, and hybrid scientific-lay forums that should mediate between the various groups (Luhmann 1993, 155; Callon *et al.* 2001; Tucker 2008; Dryzek *et al.* 2009).

### III. THE DISTRIBUTION OF RISK FROM THE EXPLOITATION OF UNCONVENTIONAL ENERGY

In what follows, I apply the theory that connects justice in the distribution of risk, risk assessments and trust in expertise and elites to explain the politics and policies that regulate unconventional energy extraction technologies in Europe. Unjust distribution of risks and rewards from new unconventional energy technologies caused heightened risk perception on the part of the victims of injustice and lowered the estimates of risk of its beneficiaries. Heightened risk perception reinforces distrust in technocratic elites that loops back to reinforce heightened risk perceptions. Consequently, protest against these new technologies in democracies often led to banning rather than regulating them. This lack of trust explains why Europeans who opposed unconventional technologies for the extraction of shale gas and tight oil were impervious to arguments about risk management and strict regulations.

Unconventional energy technologies extract shale gas, tight oil, and coal bed methane from shale rock. They include hydraulic fracturing, directional, mostly horizontal, drilling, and seismic imaging. Hydraulic Fracturing<sup>1</sup> uses industrial quantities of water mixed with sand and chemicals to fracture shale formations and release natural gas and other hydrocarbon fuels to migrate to the surface. These technologies were developed and used initially in Texas by George Mitchell, and are responsible

to a large extent for the decline in the price of natural gas and oil during the second decade of the 21<sup>st</sup> century (Tucker 2014b). The main risks in the exploration and exploitation of unconventional energy resources include wellbore integrity – if the steel and concrete well-casing fails, liquids may ascend from the depth and contaminate groundwater with brine, the dried ancient seas where the natural gas was produced that are very salty and may include naturally occurring radioactive substances, methane, and fracturing fluids. Flowback ‘produced waters’ come to the surface after properly functioning hydraulic fracturing. These are composed of fracturing fluid mixed with brine from the shale formation. The mixture can be stored on the surface and reused for fracturing. The risk is that it may not be properly stored, e.g. in open pools that may leak, and spill to contaminate the soil and the groundwater from above (rather than from below). Reusing these fluids requires the separation of solids from the liquid. If improperly disposed of, these solids (salts, chemicals, and even naturally occurring radioactive substances) can be dangerous. Another method for disposing of produced waters is by injecting them back underground, injections that have been associated with seismic events, usually less powerful than earthquakes (1-2 on the Richter scale), but sometimes noticeable. The reason for these seismic events may be that produced waters function as ‘lubricants’ when injected into small geological fault lines. Faulty valves and pipelines in natural gas wells, unconventional as well as conventional, may generate ‘fugitive’ methane emissions. As a greenhouse gas, methane is worse than carbon dioxide (IRGC 2013).<sup>2</sup>

In addition to the above uncertain *risky* externalities that can have different assessments, unconventional energy extraction has certain *cost* externalities. These costs include pressure on infrastructure and its degradation from the transportation of heavy equipment and large quantities of water and chemicals, noise, and aesthetic damage to the landscape from drilling pads. In areas of natural beauty where tourism is a mainstay of the local economy and where city folks own holiday houses,

the eyesore of drilling pads may harm the local economy, reduce property values, and the intrinsic value of holiday houses. In areas that are not tourist attractions, such costs are lower. For example, there was greater resistance to unconventional energy exploration on Poland's Baltic coast than in the Lublin basin because the Baltic coast is a summer tourist destination, while rural south-east Poland does not attract many tourists and has already borne the environmental brunt of Communist era industrialization. In areas where water for agriculture is scarce – such as Texas, Australia, and South Africa – industrial use of water may increase its cost, for example, for water-intensive rice farming in Texas.

The universal environmental benefits from developing unconventional energy include lower carbon emissions from burning natural gas in comparison with burning coal (about a third per unit of produced energy) and oil (about a half per unit of produced energy), without the safety risks associated with carbon neutral nuclear energy and the higher costs and limited supply of environmentally friendly solar and wind energy. Benefits for nation-states include the promise of energy security following reduction or even elimination of energy imports. This is particularly significant for countries that are dependent on imports from less than friendly sources. The elasticity of the price of natural gas in each country is affected by the monopoly power of the supplier (Bros 2012). Countries that have diverse sources of supply will see little change in the price of natural gas as a result of increased, unconventionally produced, local supplies, unless they are massive (Carrington 2013). But countries that are dependent on Russian supplies in East and Central Europe would be able to pay much less for their imported natural gas if they diversify their sources to become independent of the near monopoly Russian supplier.

Obviously, companies that specialize in unconventional energy exploration and exploitation stand to make a profit, otherwise they would refrain from investing or divest from disappointing projects, as they did

in many parts of Europe following the sharp decline in energy prices in the middle of the second decade of this century. Economic development in the energy sector creates jobs directly and indirectly. Above all, owners of mineral rights stand to gain royalties. In Europe, these owners are the states. When the state is run by less than entirely honest politicians and civil servants, they may derive some personal gains from selling licenses and take a cut from royalties.

#### IV. DISTRIBUTIVE INJUSTICE: RISKS, COSTS AND REWARDS

The exact environmental and economic risks and costs associated with the exploration, development and exploitation of unconventional energy resources differ from location to location, as geological formations differ from each other. With the exception of methane emissions that accelerate global warming, however, the risks associated with unconventional energy (seismic tremors, pollution, contamination of the soil and the water table from the faulty disposals of flow back waters and the transportation of chemicals) and its costs (noise and traffic, stress on the infrastructure, aesthetically displeasing landscape and so on) are *local*, in the region where the drilling takes place. In Europe, most of the rewards for the technological risks tend to be split between the companies that operate the drilling rigs and the owner of the mineral rights, namely the state, politicians and civil servants.

Historically, mineral rights in Europe were private until the 1930s. This facilitated the Industrial Revolutions of the 19<sup>th</sup> century, between the French Revolution and the First World War. Governments nationalized mineral rights and the energy sectors in Europe following the economic depression of the 1930s and then the militarization of resources during the Second World, mostly by the occupying Germans. After the Second World War, however, Communist, Christian Democrat, and Social Democrat governments did not restore natural resources to their pre-war owners. For example, ownership of oil and

gas within the land area of Great Britain was vested in the Crown by the 1934 Petroleum (Production) Act. In the subsequent 1998 Petroleum Act, ‘The Crown’ represented by the Secretary of State for the Department of Energy and Climate Change retained all sub-surface rights to hydrocarbons, such as Shale Gas. As a result, landowners do not benefit directly from natural resources like gas and oil underneath their property. The UK government grants licenses for hydrocarbon exploration and exploitation. Companies need only to gain access rights from landowners. If landowners do not grant access to their land, the law “enables the Secretary of State to confer, any right to enter on or interfere with land.” The nationalization of natural resources is a necessary condition for the unjust distribution of risks and rewards between locals who bear the external risks and costs of extractions and the political centre that reaps the rewards. Yet this is an unintended and unforeseen consequence. For a hundred years or so, Europeans have assumed that all the resources under European soil have already been discovered. This assumption had not been seriously challenged until the discovery of new but risky technologies for the exploitation of shale resources. Most of the significant discoveries before the introduction of technologies for the exploitation of unconventional energy resources were off shore. Undersea drilling did not require the cooperation of locals, although the risk of pollution from undersea accidents had to be born exclusively by locals who made their living from the sea or resided at the seashore.

This unjust distribution of risks and rewards has created a conflict in Europe between locals and local authorities who would have to assume most of the risks and costs, and elites in the centres of political and economic power who stand to gain the rewards. Consequently, the political stage is set for struggles between locals at the affected areas who perceive injustice, distrust the political centre, and overestimate the risks, and elites at the centre who underestimate risk, trust expertise, and consider locals irrational. This political conflict over the just distri-

bution of risks and rewards can be exacerbated when there are identity (ethnic, national, class, political or other) differences between the dominant group in the political centre and locals where the shale formations are concentrated.

## V. JUSTICE, POLITICS AND POLICIES

Since the controversial technologies associated with hydraulic fracturing were imported to Europe in mature form from the US where they had been invented, tested, and refined, European countries were able to choose between accepting, banning, or regulating them. There has not been a Europe-wide ‘unconventional energy’ policy. As Wyciszkievicz observes, the dispute within the European Union has not been just about NIMBY (not in my backyard) but also about whether there is one backyard, or twenty-seven backyards (2013, 112). The European Commission and its three relevant Directorates – for Energy, Climate, and the Environment – have adopted a wait and see approach, balancing pressures from member states with interests in energy security against pressure from environmentalist groups and Green members of the European Parliament to impose moratoria on the use of hydraulic fracturing. Non-binding resolutions adopted by the European Parliament have rejected more radical declarations by both sides, while the European Commission has adopted a set of non-binding recommendations.<sup>3</sup> Representatives of nations supporting and opposing the technologies and parties along the Green *versus* Conservative pro-business continuum tend to balance each other (Wyciszkievicz 2013). European regulations about the extraction of minerals that had been enacted before the introduction of unconventional energy technologies have been interpreted for the new technologies (Musialski 2013). Regulatory decisions about the new unconventional energy technologies have been delegated to member states. European policies and laws about unconventional energy extraction vary from a total ban on hydraulic fracturing in

France, through moratoria in Bulgaria and the Czech Republic, to enthusiastic support in Poland<sup>4</sup>:

| Countries                                                                                                              | Policies                                                   |
|------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|
| France                                                                                                                 | Banned exploitation, but not experimentation               |
| Bulgaria, Luxemburg                                                                                                    | Indefinite moratoria                                       |
| Czech Republic                                                                                                         | Temporary moratorium                                       |
| Germany                                                                                                                | No federal policy                                          |
| Baden-Wittenberg                                                                                                       | No licenses                                                |
| Lower Saxony                                                                                                           | Granted licenses                                           |
| Ireland                                                                                                                | Contested                                                  |
| Norway, Belarus, Moldova, Croatia, Bosnia, Belgium                                                                     | Allowed in principle, but no licenses have been issued yet |
| Romania                                                                                                                | Cycles of moratoria and licensing                          |
| Ukraine, Lithuania, Estonia, Hungary, Serbia, Slovenia, Sweden, Denmark, Netherlands, Spain, Portugal, Greece, Turkey. | Limited issuance of licenses                               |
| Poland                                                                                                                 | Enthusiastic government and popular support                |
| United Kingdom                                                                                                         | Initial moratorium, followed by strong government support  |

It is striking that while US states, regions, and municipalities have introduced varying types and levels of regulations (Robeck & Bennett 2013, 54-57), European states have either banned unconventional technologies such as hydraulic fracturing altogether<sup>5</sup>, put a moratorium on them, *de facto* banned them through administrative means by issuing no licenses for drilling or very few licenses for drilling only in sparsely

populated areas, issued licenses for exploration but not for exploitation, or at the other extreme allowed or even encouraged drilling.

Since regulatory frameworks develop with technologies, late adopters of unconventional technologies were able to draw regulatory conclusions from what US states have learned from their mistakes and experience with unconventional energy (IRGC 2013, 45-59). European countries received regulatory advice from the US State Department's Unconventional Gas Technical Engagement Program (UGTEP) (Tucker 2014b, 90-91). Liability regulations may ensure that if risks are realized, the costs would not be externalized to the residents of affected areas but be borne by the responsible parties. The water table could be protected by regulating the depth of hydraulic fracturing and the integrity of the wellbore and well-casing. Regulations could mandate the publication of the chemicals that are used in hydraulic fracturing, or the addition of a harmless unnatural trace material, and the measuring of the baseline chemical composition of water sources. If this composition changes after hydraulic fracturing, it should be easy to prove the liability of the companies. Regulations can minimize risks from storing, treating and disposing produced waters. Regulations can also minimize the risk from injecting flowback waters into the ground by prohibiting injections into geological fault lines and limiting their strength and quantity. Mandated better valves and designs of production, transmission and distribution of natural gas, and regular maintenance and inspection of equipment can minimize methane emissions (IRGC 2013; Task Force on Shale Gas 2015).

The absence of regulations and regulatory debates reflect the unjust distribution of risks without rewards to locals who then protest the very use of the risky technologies. Risk bearing locals have no interest in safety regulations to mitigate the risks because they do not want to have to bear any risks with no prospective rewards. When political protest against injustice is successful it leads to banning of the technology either *de jure* through legislation or *de facto* by not issuing licenses. If locals lose the struggle, the technologies remain unregulated because the elites who

stand to benefit from the unjust distribution of risks have no interest in mitigating them.

## VI. UNACCOUNTABLE TECHNOCRATIC FAILURES AND DISTRUST

In this and the following section, I highlight two other factors that affect levels of trust in elite technological risk management: risk management failures of unaccountable elites and geopolitical threats.

Giddens (1990, 136) called the attitude of experts who believe in unbridled enlightenment rationality ‘sustained optimism’. Until the 1990s, in countries like France and the United Kingdom, policies about science and technology were decided by technocratic experts, civil servants with sustained optimism. Technocratic elites can plan for a longer term than politicians whose temporal horizons proceed from one election to the next. Technocrats can also be better informed and more dispassionate about risk assessment and management than voters. However, when cohesive technocratic elites make a mistake and mismanage risk, they find it very difficult to admit responsibility and correct themselves, tending rather to deny, obfuscate, and close ranks. Consequently, it loses the public’s trust in their ability to manage science and technology risks (Jasanoff 2005). The recent failures of technocratic elites in the private and public sectors to pre-empt and then successfully manage the fallout from the recession of 2007-8 have eroded trust in their technocratic skills and partly led to populist revolts against them, driven by unleashed passions, resentments and fears. But already towards the end of the previous century, failures in managing the Mad Cow disease outbreak in the UK and the HIV contaminated blood crisis in France eroded trust. In some Western European societies, rifts have emerged between what Dryzek *et al.* (2009) called ‘promethean elites’ and ‘precautionary publics’. European technocratic elites have maintained a promethean enlightenment ethos of belief in science and technology. But their precautionary populations have come to distrust their zeal. The legal and political result of this distrust

has been the Precautionary Principle that among other provisions considers new technologies risky until proven otherwise (Callon *et al.* 2001).<sup>6</sup> The French rejection of hydraulic fracturing can be traced back to the adoption of the Precautionary Principle as part of French law that predated the prospects for development of unconventional energy industry in France (Inchauspé 2013; Beytout 2013). Environmentalist groups and green politicians and parties represent precautionary publics.

By the time hydraulic fracturing arrived in France, the cleavage between technocratic elites and distrusting masses had already formed. The distrust caused a reinforcing causal feedback loop when the French technocratic elite did not attempt to convince the citizens of its low risk assessment of hydraulic fracturing but took popular distrust and opposition as a given and attempted to smuggle in its optimistic promethean agenda by stealth, namely by granting drilling concessions before the public noticed. Technocrats attempted to use the same methodology earlier to decide without accountability where and how to store nuclear waste and to admit genetically modified food, when the public was distracted or away on vacation, with similar results. When the ‘plot’ was uncovered it caused public outrage, and deepened further distrust in unaccountable technocratic elites (Callon *et al.* 2001). Consequently, there is a reinforcing causal feedback loop. Distrust in technocratic risk management leads to stealthy unaccountable decisions about risk assessment and its distribution, which further deepen the distrust and so on. ‘Activist farmer’ José Bové, a Green Party Euro MP who began his career by flattening a McDonald’s outlet, observed at the beginning of 2011: “Opposition is building because people are shocked by the way the state pushed through drilling permits without any debate. It shows disdain for the population and elected representatives” (Patel 2011). On March 16<sup>th</sup> 2011, French Industry Minister Eric Besson stated that while chemicals used in fracturing “[...] have caused considerable damage in the U.S. and Canada [...] some industry representatives say that there may be clean technology that would permit production of shale gas without causing what we have seen

in the U.S.” (Patel 2011). But after suffering a crushing defeat in regional elections in March 2011, President Sarkozy’s ruling UMP party drafted a bill to ban hydraulic fracturing in France. The law did not prohibit shale gas exploration and technological experimentation, only economic exploitation. The strategy may have been to get the issue out of the public arena until after the presidential election, while allowing companies and scientists to continue exploring and preparing for actual exploitation (Pilgrim 2011; Tomasi and Nicolet 2013). Following the victory of President Hollande in the presidential elections and the establishment of a Red (Socialist Party)-Green coalition, the law was not revised. The French political cleavage is not between left and right, but between a promethean elite and precautionary public and populist politicians (Dryzek *et al.* 2009). Within the then ruling Socialist Party some politicians wanted to reopen the issue, and it seemed they had a chance in the summer of 2013 in the context of a review of France’s energy policy, after the president dismissed the Environment Minister, a strong opponent of unconventional energy. Nevertheless, President Hollande put an end to the debate in his 2013 Bastille Day speech when he declared that there will be no shale gas exploitation in France while he was president.

In post-communist countries in Europe, distrust in unaccountable political and administrative elites has deeper historical roots than in France or the UK. In countries that were totalitarian or are kleptocratic, political and economic elites have a legacy of disregarding the legitimate interests and rights of their citizens, including their right to live in a reasonably safe environment. Communism prioritized economic growth, ‘progress’, and industrialization at all costs, largely ignoring environmental concerns. In the absence of institutions that could protect the environment and citizens’ rights, Communist regimes habitually devastated the environment and ignored the effects of pollution on public health. While new institutions of the rule of law, independent judiciaries, and even environmental protection agencies were created, they have been considerably weaker than in established liberal democracies, and sometimes

amount to mere facades or ‘Potemkin villages’. In many post-Communist countries, politicians have not expected to stay long in politics, so they have attempted to get rich quickly from their office by whatever means necessary, legal and otherwise (Tucker 2015). Consequently, citizens of post-Communist countries have not trusted their political elites not to sell environmental or technological risks to the highest bidder, and feared the consequences of decisions they make about new technologies. According to Transparency International, for example, Bulgaria is the second most corrupt country among the 28 EU member states. Chevron allegedly paid 30 million euros for a concession to drill for unconventional energy in Bulgaria. This money has not been accounted for. Demonstrations first caused a moratorium on explorations of unconventional energy. Then a second set of demonstrations in 2013 protesting low living standards and high utility bills in the poorest member of the European Union brought down the Bulgarian government (Goussev 2014).<sup>7</sup>

## VII. THE GEOPOLITICS OF TRUST

A widely shared sense of threat from external enemies may countervail distrust in national elites and politicians to lower popular risk estimates. In comparison with the geopolitical risk posed by Russia, the technological risk of unconventional energy may not be as high as it had seemed. As a legacy from their history, countries in the former Soviet Bloc are much more dependent on imported Russian energy than Western European countries with diversified foreign suppliers. Russia has been using its monopoly power to hike prices (Bros 2012) and to pressure countries to toe its political line and serve its geopolitical interests, for example by agreeing to allow prospective Russian pipelines (Southstream) to pass through their territory. Russia uses ‘carrots’ in the form of price discounts to compliant countries and ‘sticks’ in the form of cutting supplies in the winter to countries, like Ukraine, that defy it. Local energy resources reduce the monopoly power of suppliers, even when countries must still

import some or much of their energy. European technocratic elites understand well the elasticity of energy prices, especially natural gas, though consumers in Eastern Europe do not necessarily understand the link between Russian monopoly power and the price they pay for their utilities (Tucker 2014b). Still, Eastern Europeans are acutely aware of the risk posed by dependence on Russian supplies. This threat trumps protests against unjust risk distribution and distrust of political elites. Poland, the Baltic States, and Ukraine have been acutely reminded of their dependence on Russian energy supplies when supplies to Europe were suspended during disputes between Russia and Ukraine (Pirani *et al.* 2009). There was thus a significant gap in Poland between the popularity of unconventional energy and trust in politicians. A 2011 survey found that 73% of Poles supported the exploration and development of shale gas resources, while only 4% opposed it (Kahr *et al.* 2012). At the same time, only 43% of Poles trusted their government, a figure that fell to 28% in 2012 (Edelman 2014). Shale exploration even appeared popular on the local level. For example, the mayor of Lesniowice, where an initial exploratory well was drilled, stated: “It would create more competitiveness on the gas market, which is now dominated by Russia, and one side would not be able to force anything unilaterally anymore” (Kahr *et al.* 2012). Kenarov (2012) observes that “[T]he real reason shale gas has made such a stir in Poland is buried deep in the emotional sediments of history, where Russia embodies the perpetual nemesis, the perpetrator of the greatest crimes against the Polish nation. The hope is that shale gas would reduce Poland’s reliance on Russian fossil fuels, while bringing the country closer to its ally the United States, a pioneer of shale gas extraction.” There have also been local protests against hydraulic fracturing in Poland, but they were limited to some of the rural communities where drilling should have taken place. As one farmer observed, “I think it is just impossible to do it here. We have no interest in it; we don’t want it because the gas will go somewhere else. Someone else will have profits; we will have no advantages from it, no positive results.” Another farmer claimed that

“Chevron came here like John Wayne into a saloon. They just organized a meeting a week before they wanted to start the ground works. They first signed the agreements and then informed people what was going to happen here.” By contrast, the farmer who leased his own topsoil for the drill pad said “I don’t work with this industry, so I don’t know details. I don’t know if this is good or bad. I believe if the government is giving them concessions and allowances to drill [then] those people know what they are doing, and that is why they are sitting there in all those ministries and offices, and they are supposed to know whether this is a safe thing or not. How can I know such a thing? If they receive permits, then I know that this is something that is not harmful. ... I still talk to people like I used to. We are still neighbors, but I don’t believe in all they say because we were together in the municipality and the district offices, and the officials there said perhaps there is some other company that is paying them [the upset neighbors], perhaps it is Gazprom, a Russian company that is paying them because they [Gazprom] don’t want this drilling to be going on here” (Beras 2012). Since most Poles are urban and anxious about their dependency on Russian energy, they voted for governments that would impose, if necessary, the development of an energy sector on reluctant local communities (Materka 2013).

In comparison, many Bulgarians do not feel threatened by their dependence on Russian energy (Goussev 2014; Jirušek *et al.* 2017; Smillov 2018). Bulgarians have been traditionally friendly to Russia with whom they shared the Orthodox faith and the Ottoman enemy. Without a sense of geopolitical risk from dependence on Russian energy to balance unjust distribution of risks and rewards and distrust of politicians, Bulgaria imposed a moratorium on shale gas explorations.

### *Comparison with the US*

Several positive European cases confirm the theoretical model that links unjust distribution of risks and rewards from the exploitation of

unconventional energy with distrust in technocratic elites and high-risk assessments. Another way of demonstrating the theory is by using J. S. Mill's method of difference, i.e. find cases where the cause – the unjust distribution of risks and rewards – is absent or reduced and examine how the effects differ. In the US, land owners usually own mineral rights down to the centre of the earth. They and local authorities that tax their income therefore receive much higher rewards than in Europe if they choose to accept the risks involved in the exploitation of minerals. In the US, local lower risk assessments result from commensurability between the choice to accept risk and rewards in the form of regular royalties. Royalties correlate the distribution of risks with the distribution of rewards and internalize the cost of risk for those who accept it. The periodic regularity of royalty payments and the fulfilment of expectations over a period of time can build trust. The advanced US legal system of liability laws and litigation and popular knowledge thereof through movies like *Erin Brockovich* have built some trust in the legal and judiciary systems that may mitigate distrust in corporations and experts.

The legal concept of property rights as a bundle of properties that can be sold, retained, or given away (cf. Honoré 1961) allows individuals and corporations in parts of the US to sell land and houses while retaining mineral rights for what lies beneath them, especially when home and land buyers do not read the fine print about mineral rights or decide its implications too implausible to be taken seriously. Pennsylvania even separated mineral rights from rights for oil and gas; each can be held or sold separately. Consequently, unjust distributions of risk can emerge when owners of the topsoil are denied the choice of assuming or refusing risk from unconventional energy exploitation and receive no rewards. Distrust of corporations that reap the rewards, and high estimates of risk follow, just as the European theoretical model would predict. US citizens whose land is atop discovered mineral resources they do not own perceive risks in new drilling technologies and oppose the exploitation of mineral resources. It does not matter if the royalties

accrue to European states or US corporations, owners of the top soil perceive risks that owners of mineral rights who choose to accept risks for rewards do not. In Colorado, for example, *Oakwood Homes* sold homes it built but retained the mineral rights which it then leased to Anadarko Energy, which in turn leased them to ConocoPhillips. The home owners were forced to bear the risks and costs associated with drilling under their homes with no rewards. They then formed *Green Valley Ranch Families against Fracking*, which believes that “fracking is hurting the environment” (Conlin and Grow 2013; Ortner 2012). Similar unjust distribution of risks that result from the separation of land ownership from ownership of mineral rights occurred also in Texas, North Carolina, and Florida (Conlin and Grow 2013).

### *For a Distributive Justice*

It is possible to attempt to reduce risk perceptions and distrust in Europe by introducing a more just distribution of risks and rewards, by sharing mineral royalties with locals, as was attempted to a limited degree in the UK.

The British Conservative-Liberal coalition government of David Cameron identified shale gas and oil as possible substitutes for the UK’s depleted North Sea gas and oil fields. In an era of economic austerity and low growth following the recession, the prospect of developing a new industry with new jobs and economic development was alluring. The government encouraged the creation and development of a shale energy industry through tax breaks for explorers and producers. “The Treasury has set a 30% tax rate for onshore shale gas production. That compares with a top rate of 62% on new North Sea oil operations and up to 81% for older offshore fields.” Chancellor George Osborne supported this new tax regime to “[...] make [it] the most generous for shale in the world” because shale gas “[...] has the potential to create thousands of jobs and keep energy bills low for millions of people” (Macalister and

Harvey 2013). The government further facilitated the licensing process, fast tracking the licensing of wells by taking away decision power from local councils and centralizing it with the government when the councils did not license drilling for shale within sixteen weeks (Spalding and Murphy 2015). Though the mineral rights and royalties belonged to the Crown, British courts recognized the rights of free hold landowners to their land down to the centre of the earth. Drilling below their land was therefore recognized in 2010 as trespass. This led to protracted negotiations between companies and landowners over payments. To circumvent this, the Infrastructure Act of 2015 (which holds in England and Wales, but not in Scotland and Northern Ireland) permits drilling without permission below 300 meters under private land, thereby allowing horizontal drilling under privately owned land without the permission of the owners. Payment for this use is left to be determined by the Secretary of State. Thus far, such payments have been voluntary.

Shale exploration in England led then to predictable local resistance from landowners without mineral rights, for example, in Balcolme village in West Sussex where Cuadrilla was forced to halt its operations (Community Evidence 2014). Facing local resistance, the British government attempted to ‘sweeten’ exploration for shale gas by redistributing some of the rewards from the state to the local residents. Communities that allowed exploration of unconventional gas and oil were promised “£100,000 in community benefits per well-site.” If those sites turned into producing sites, communities would receive “1% of revenues” from sold commodities, which was estimated to “provide £5-10m per well, spread over 25 years, but mostly in the first 10” (Department of Energy 2013). This ‘sweetener’ was grossly insufficient because the rewards for those forced to assume risk were far too low and deliberately vague. There were no concrete specifications of the meaning of ‘communities’. They could mean parish councils, the lowest level of local government in the UK, or district or county councils that administer larger areas. Tessa Munt, a Liberal Democrat, argued that the district and county councils may not

even spend the sweetener among those forced to assume risks in the affected areas. The government did not effectively persuade individual citizens of the benefits of unconventional energy for their economic livelihood (Munt 2014). Nevertheless, from the perspective of our model, this is a step in the right direction; it needs to go further toward offering locals just returns for choosing to assume risk.

## VIII. CONCLUSION

When risks and rewards from new technologies are distributed unjustly, involuntarily, and without correspondence between risks and rewards, the victims of the unjust distribution of risks who are forced to pay the cost of externalities resist it, distrust the technocratic elites and experts who make decisions about risk distribution and management and assess the risks as more severe than technocratic elites and experts. Justice in the distribution of risk and corresponding rewards is not only a philosophical abstraction and ethical theory, it is also a political force. It has a social explanatory as well as normative force.

Our application of Beck's theory about risk distribution, Luhmann's and Giddens' theories of risk assessment, and Schmidtz' theory of justice to understanding the social and political reaction to new 'unconventional' technologies for extracting fossil fuels from shale rock discovered correlations between justice in the distribution of risks and rewards and risk assessment and trust, reinforced by a feedback causal loop whereby high assessment of risk and distrust reinforce each other. The unaccountability of technocratic and political elites increases distrust and generates a feedback causal loop that increases the unaccountability of elite decision making, when elites adjust to popular distrust by attempting to make decisions about the distribution of risk from new technologies in stealth or when the population is distracted. Geopolitical risk may balance technological risk assessment when assuming technological risks reduces geopolitical risks.

Schematically, the causal theoretical model looks like this:



## WORKS CITED

- Balcombe Residents Association. 2014. "Community Evidence to the Lords Select Committee on the Economic Impact on UK Energy Policy of Shale Gas and Oil." January 21. <http://www.frackfreebalcombe.co.uk/page7.php> [accessed August 25, 2018].
- Beck, Ulrich. 1992. *Risk Society: Towards a New Modernity*. Translated by Mark Ritter. London: Sage.
- Beras, Erika. 2012. "Zurawlow and Chevron." WESA Pittsburgh's NPR News Station. <http://wesa.fm/post/zurawlow-and-chevron> [accessed August 25, 2018].
- Beytout, Nicolas. 2013. "Principe de démagogie." *L'Opinion: Quotidien liberal, politique, économie, international*, July 1.
- Bros, Thierry. 2012. *After the US Shale Gas Revolution*. Paris: Editions Technip.
- Callon, Michel, Pierre Lascoumes and Yannick Barthe. 2001. *Agir dans un monde incertain: Essai sur la démocratie technique*. Paris: Seuil.
- Carrington, Damian. 2013. "Lord Browne: Fracking Will Not Reduce UK Gas Prices." *The Guardian*, November 28. <http://www.theguardian.com/environment/2013/nov/29/browne-fracking-not-reduce-uk-gas-prices-shale-energy-bills> [accessed August 25, 2018].
- Conlin, Michelle, and Brian Grow. 2013. "Across the U.S., Thousands of Families Are Moving into New Houses Where, Amid an Unprecedented Energy Boom, Their Builders or Developers Have Kept the Mineral Rights for Themselves." Reuters, October. <http://graphics.thomsonreuters.com/13/10/USA-FRACKING:RIGHTS.pdf>.
- Department of Energy & Climate Change. 2013. "Developing Onshore Shale Gas and Oil: Facts about 'Fracking'." <https://www.gov.uk/government/uploads/system/>

- uploads/attachment\_data/file/265972/Developing\_Onshore\_Shale\_Gas\_and\_Oil\_Facts\_about\_Fracking\_131213.pdf.
- Dryzek John, Robert Goodin, Bernard Reber and Aviezer Tucker. 2009. “Promethean Elites Meet Precautionary Publics: The Case of GM Foods.” *Science, Technology & Human Values* 34/3: 263-288.
- Editorial. 2013. “Frack to the Future: Unconventional Gas In Europe.” *The Economist*, 2 February. <https://www.economist.com/business/2013/02/02/frack-to-the-future> [accessed August 25, 2018].
- Goussev Viktor Nikolaev, Simon Devey, Bianca Schwarzenburg, Marco Althaus. 2014. “Shale Gas U-Turns in Bulgaria and Romania: The Turbulent Politics of Energy and Protest.” *European Managements and Public Affairs* 1/2.
- Giddens, Anthony. 1990. *The Consequences of Modernity*. Stanford, CA: Stanford University Press.
- Gillet, Kit. 2013. “Bulgarian Students Lead Wave of Protest.” *The Guardian*, 26 November. <http://www.theguardian.com/world/2013/nov/26/bulgaria-student-protest-corruption> [accessed August 25].
- Graham, John, John A. Rupp, Austin Mitchell. 2013. *Risk Governance Guidelines for Unconventional Gas Development*. Lausanne: International Risk Governance Council.
- Honoré, A. 1961. “Ownership.” In *Oxford Essays in Jurisprudence*. Edited by A. Guest, 107-147. Oxford: Oxford University Press.
- Inchauspé, Irène. 2013. “Ce principe de precaution qui paralyse la France.” *L’Opinion: Quotidien liberal, politique, économie, international*, 1 July.
- Jasanoff, Sheila. 2005. *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton: Princeton University Press.
- Jirušek, Martin, Tomáš Vlček and James Henderson. 2017. “Russia’s Energy Relations in South-eastern Europe: An Analysis of Motives in Bulgaria and Greece.” *Post-Soviet Affairs* 33/5: 335-355.
- Kahr, Michael Braden, Braden Reddall and Gabriela Baczyńska. 2012. “Insight: Poland’s Shale Gas Play Takes on Russian Power.” *Reuters*, 9 February. <http://www.reuters.com/article/2012/02/09/us-poland-shalegas-idUSTRE8180PM20120209> [accessed August 25, 2018].
- Kenarov, Dimiter. 2012. “Poland’s Shale Gas Dream.” *Pulitzer Center on Crisis Reporting*, 29 December. <http://pulitzercenter.org/reporting/poland-shale-gas-industry-development-energy-security-economy-environment> [accessed August 25, 2018].
- Kriebel, David, Joel Tickner, Paul Epstein, John Lemons, Richard Levins, Edward L. Loechler, Margaret Quinn, Ruthann Rudel, Ted Schettler, Michael Stoto. 2001. “The Precautionary Principle in Environmental Science.” *Environmental Health Perspectives* 109/9: 871-876.
- Luhmann, Niklas. 1993. *Risk: A Sociological Theory*. Translated by Rhodes Barrett. Berlin: Walter de Gruyter.
- Macalister, Terry, and Fiona Harvey. 2013. “George Osborne Unveils ‘Most Generous Tax Breaks in World’ for Fracking.” *The Guardian*, 18 July. <http://www.theguardian.com/>

- politics/2013/jul/19/george-osborne-tax-break-fracking-shale-environment [accessed August 15, 2018].
- Materka, Edyta. 2013. "Handling Public Discontent: Drawing Some Lessons from Pomerania Turmoil in Local Communities." In *Shale Gas in Europe: A Multidisciplinary Analysis with a Focus on European Specificities*. Edited by Cécile Musialski *et al.*, 393-416. Deventer: Claeys & Casteels.
- Munt, Tessa. 2014. "Why Fracking Incentives Should Go to Parish Councils - Even If They Are Worth Millions." *Mirror Online*, 21 January. <http://www.mirror.co.uk/news/uk-news/tessa-munt-fracking-incentives-should-3043328> [accessed August 25, 2018].
- Musialski, Cécile. 2013. "The EU Legal and Regulatory Framework as Currently Applicable to Shale Gas Extraction and Commercialization." In *Shale Gas in Europe: A Multidisciplinary Analysis with a Focus on European Specificities*. Edited by Cécile Musialski *et al.*, 61-110. Deventer: Claeys & Casteels.
- Ortner, Richard. 2012. "Anti-Fracking Group Hopes New York Times Ad Wins First Lady's Support." *Fox 31 Denver*, 13 May. <http://kdvr.com/2012/05/13/colo-anti-fracking-group-hopes-new-york-times-ad-wins-first-ladys-support/> [accessed August 25, 2018].
- Patel, Tara. 2011. "The French Public Says No to 'Le Fracking'" *Bloomberg Business Week*, 31 March. [http://www.businessweek.com/magazine/content/11\\_15/b4223060759263.htm](http://www.businessweek.com/magazine/content/11_15/b4223060759263.htm). [accessed August 25, 2018].
- Pirani, Simon, Jonathan Stern, and Katja Yafimava. 2009. *The Russo-Ukrainian Gas Dispute of January 2009: A Comprehensive Assessment*. Oxford: Oxford Institute for Energy Studies.
- Robeck, Mark R. and Bennett Michael. 2013. "Shale Gas in the United States: An Institutional Comparison." In *Shale Gas in Europe: A Multidisciplinary Analysis with a Focus on European Specificities*. Cécile Musialski *et al.* 35-60. Deventer: Claeys & Casteels.
- Schmidtz, David. 2006. *Elements of Justice*. Cambridge: Cambridge University Press.
- Smilov, Daniel. 2018. "Bulgaria Torn Between Russia and the West." *Deutsche Welle*, 31 May. <https://www.dw.com/en/bulgaria-torn-between-russia-and-the-west/a-44027331>. [accessed August 25, 2018].
- Spalding, Darren & Murphy Eimear. 2015. "UK Government Continues the Fight for Shale." *Natural Gas Europe*, 6 October. <http://www.naturalgaseurope.com/uk-government-continues-the-fight-for-shale-25961>. [accessed August 25, 2018].
- Task Force on Shale Gas. 2015. *Second Interim Report: Assessing the Impact of Shale Gas on the Local Environment and Health*. London: Task Force on Shale Gas.
- Tomasi, Thierry and Elsa Nicolet. 2013. "France: A Step-Back: The French Ban and its Aftermath." In *Shale Gas in Europe: A Multidisciplinary Analysis with a Focus on European Specificities*. Edited by Cécile Musialski *et al.*, 451-464. Deventer: Claeys & Casteels.
- Tucker, Aviezer. 2008. "Pre-emptive Democracy: Oligarchic Tendencies in Deliberative Democracy." *Political Studies* 56: 127-147.

- Tucker, Aviezer. 2014a. “Epistemology as a Social Science: Applying the Neuman-Rubin Method to Explain Expert Beliefs.” In *Experts and Consensus in Social Science*. Edited by Carlo Martini and Marcel Boumans, 155-170. Dordrecht: Springer.
- Tucker, Aviezer. 2014b. “The New Energy Force Field.” *The American Interest* 9/5: 85-95.
- Tucker, Aviezer. 2015. *The Legacies of Totalitarianism: A Theoretical Framework*. New York: Cambridge University Press.
- Wyciskiewicz, Ernest. 2013. “European Institutions in the Debate on Shale Gas.” In *Shale Gas in Europe: A Multidisciplinary Analysis with a Focus on European Specificities*. Edited by Cécile Musialski *et al.*, 111-130. Deventer: Claeys & Casteels.

## NOTES

1. The use of ‘fracking’ to refer to hydraulic fracturing originated with protesters against the technology who added the letter *k* to a derivative of ‘fracturing’. I leave it to the reader to figure out the desired associations. I use the technical, family friendly, and politically neutral ‘hydraulic fracturing’.
2. See this report and that of the Task Force on Shale Gas (2015) for bibliographic lists of sources.
3. See: [http://ec.europa.eu/environment/integration/energy/unconventional\\_en.htm](http://ec.europa.eu/environment/integration/energy/unconventional_en.htm)
4. For an overview of European policies see *The Economist* (2013). For updates, I recommend the electronic magazine *Natural Gas Europe* <http://www.naturalgaseurope.com> which combines original articles with aggregated articles from the press.
5. To take an extreme example, in Bulgaria, an overzealous parliament enacted regulations so strict that they also prohibited the small scale pre-existing Bulgarian conventional natural gas industry, requiring the parliament to scale back the regulations.
6. “The precautionary principle, proposed as a new guideline in environmental decision making, has four central components: taking preventive action in the face of uncertainty; shifting the burden of proof to the proponents of an activity; exploring a wide range of alternatives to possibly harmful actions; and increasing public participation in decision making” (Kriebel *et al.* 2001, 871).
7. In November 2013, a student protest movement “[...] rapidly escalated into the biggest rolling wave of demonstrations since the collapse of communism 24 years ago. [...] Fury at corruption and nepotism, youth unemployment at 28.7%, low wages and limited job opportunities that force qualified Bulgarians to find work overseas, and a sense that those in power have for too long put their own interests above those of the country, have coalesced this year into one long pulse of anger” (Gillet 2013).

---

## BOOK REVIEWS

---

Robert L. HOLMES. *Introduction to Applied Ethics*. London: Bloomsbury, 2018. 527 pp.

Robert Holmes' *Introduction to Applied Ethics* is primarily intended as a book to be used in first-year applied ethics seminars, but it can also be used by anyone wishing to get a first glimpse of some of the major questions discussed in applied ethics. It results from a teaching experience of thirty years and has thus been tested in practice. Written for a large public, the book does not require any deep preliminary knowledge of philosophy, and the knowledge required to tackle the problems and questions presented in the book is given by the author in the context of his discussions. Thus, in each chapter we find several "Definition Boxes" – defining the key notions of the chapter, e.g. racism or war – and "Theory Boxes" – introducing key metaethical distinctions, e.g. the distinction between descriptive and emotive meaning, the basic tenets of ethical theories, e.g. utilitarianism, general philosophical notions, e.g. the distinction between a valid and a sound argument, but also more general information, e.g. the distinction between a zygote, a morula, an embryo and a foetus. A third kind of box contains "Key Quotes", short passages, often from philosophers, which take a stand or illuminate the problem from a certain perspective. Besides these, we also find longer texts in boxes with a grey background. These texts are taken from philosophical works, official documents, newspaper articles, etc. Their main function is to allow a discussion to emerge. At the end of each chapter we find "Study questions" – ten per chapter –, so that readers can check if they have correctly understood what precedes. To help the reader in this task, Holmes indicates where to look in the chapter in order to find the answer or solution. And for those who seek more resources, there is a reference to a link where they can be obtained.

The book is subdivided in five parts. The first part tackles the questions of "Race, Gender and Ethnicity" and is subdivided in five chapters, the first three relating directly to the three concepts indicated in the part's title, whereas chapter 4 analyses affirmative action and chapter 5 is about sexual harassment. Though the author generally tries to remain neutral, he takes a clear stand against racism and sexism, showing that neither the racist, nor the sexist can bear the burden of proof that is theirs.

As to this first part, one may wonder why the author has completely left out sexual ethics and why he has one chapter on sexism and another chapter on sexual harassment. While sexual harassment may be condemned on grounds other than sexism, the two

could have been discussed in one chapter – e.g. Gender relations –, thus leaving space for a chapter on sexual ethics.

Moreover, one may wonder why there is no chapter or part of a chapter devoted to xenophobia, which is, ethically speaking, much more interesting than racism. Whereas the racist maintains that some races are intrinsically superior to others, xenophobes are afraid of people who are different from them and refuse to have these people – which they do not have to consider as intrinsically inferior – live together with them in the same country. The question: “which degree of difference are we morally bound to accept or tolerate?” is ethically much more interesting than the questions relating to racism. Chapter 3 briefly alludes to this. Another important question to be asked in this context is “is the State morally allowed to give preference to nationals when distributing certain goods, like social help, when it is clear that not all those who need these goods can get them?”

In this context, one may also wonder why the author has limited his discussion of immigration to Hispanic/Latino immigration. While there is no doubt that this kind of immigration is important in the US, it would nevertheless have been more appropriate to discuss the question of immigration more generally and ask questions like “is a country responsible for a large part of global climate change ethically bound to accept refugees from small Pacific islands about to be submerged by the sea?” or “is a country responsible for the political chaos in Libya ethically bound to accept refugees from that country?” The late French Prime Minister Michel Rocard may have been right when saying that France could not take up all the misery in the world. But what if France has contributed to producing that misery?

The title of Part II is “Profit and the Plight of Others” and is, to sum up, devoted to questions relating to economy. The first chapter of this part – chapter 6 in the continuous chapter numeration – discusses corporate responsibility. It is followed by a chapter on poverty and world hunger and by a chapter comparing capitalism and socialism.

Part III is devoted to animals and the environment, each of these topics being discussed in one of the two chapters making up the part – beginning, contrary to what the title suggests, with a chapter on the environment.

Part IV – Autonomy and the Individual – contains three chapters. The chapter on privacy has its place in this part. But one may wonder whether the same can be said about the chapters on abortion and medical aid in dying/euthanasia. While the abortion debate in the US context is often framed in the language of privacy and individual autonomy, one should not forget that this is mainly due to the fact that when fighting for the legal possibility of abortion, one of the only possibilities one had was to refer to the constitutional amendment protecting personal autonomy. Looked at from another perspective, the question of abortion could also have been discussed in Part V – The Nonconsensual Taking of Human Life (subchapter 10 of chapter 12 discusses abortion as a case of “killing the innocent”). If not in Part I, sexual ethics would also have fitted in Part IV.

In the fifth and last part of the book, the author discusses the death penalty and, lumped together in one chapter, terrorism and war. As to the latter, Holmes states that “[...] war is almost universally accepted” (482). It would have been more prudent to say that war is almost universally tolerated, as the idea of acceptance makes it appear as a good, whereas the idea of toleration implies that one sees it as a kind of necessary evil.

Although Holmes primarily wants to present and discuss problems and not theories or authors, one may wonder whether it is appropriate to mention Aristotle only once, and that is with reference to his thesis on the moral superiority of men over women, and Kant only twice – once also with the thesis of man's moral superiority over women and once with his thesis that we are morally obliged to develop our rationality.

Despite the criticisms formulated here, it has to be said that from a pedagogical point of view, Holmes' book is highly recommendable for beginners. Judging each chapter separately and leaving out the question of whether the author should not also have discussed this or that topic, Holmes' *Introduction to Applied Ethics* is written in a clear and rigorous style, presenting and discussing the basic arguments in an exemplary way, thus making it possible to evaluate them. It could thus be said that the book is not only a useful introduction to applied ethics, but also an introduction to moral reasoning and to the analysis of philosophical arguments. The extensive glossary at the end of the book is very helpful for readers who quickly want to look up a particular notion and do not remember in which chapter of the book it was introduced in a Definition or Theory Box. One can only regret that the glossary only contains notions and not also names – these can nevertheless be found in the index.

Norbert Campagna  
Université du Luxembourg

Thomas PÖLZLER. *Moral Reality and the Empirical Sciences*. New York: Routledge, 2018. 246 pp.

Thomas Pölzler's *Moral Reality and the Empirical Sciences* examines whether social scientific research can usefully inform the traditional philosophical study of morality. This is a timely topic because there has in fact been much recent reference to scientific findings in philosophical work and even collaboration between social scientists and philosophical ethicists – indeed, with some philosophers themselves becoming so-called experimental philosophers. Pölzler narrows his focus still further to the bearing of science on the meta-ethical issue of whether morality is real, that is, whether there are objective moral facts or truths. However, Pölzler also intends that his investigation have implications for empirical/conceptual 'intertwining' (his word) more broadly.

The book is a model of clarity. Pölzler's arguments and plan of discussion are clearly laid out and adhered to from beginning to end. A great many complex issues are taken up, but the reader always knows why and what the upshot is. Needless to say, of course, this monograph will have particular appeal, in terms of both content and form, to the specialist.

Pölzler does not himself take a position on the meta-ethical issue of whether morality is real. Where Pölzler takes his stand (or stands) is on the question of whether science can help resolve the meta-ethical issue. His answer is twofold: (i) most of the empirical work to date is not helpful in resolving any of the philosophical issues it is (supposedly) relevant to, but (ii) the situation can be rectified to some degree. Thus Pölzler strives to be critical but in a constructive manner; although there is one *tour de force* after another

of subtle and insightful criticism, his overall purpose is benign toward empirical philosophizing. This book is particularly impressive in the precise way it marshals the flaws it exposes for making positive suggestions about how to improve things in future.

Nevertheless, the overall effect, at least on this reader, is quite the opposite. One gets the impression that the problems are so numerous and intractable that scientific research may not be able to contribute to resolving moral objectivity after all. (I am reminded of how Daniel Kahneman in *Thinking, Fast and Slow* [Farrar, Straus and Giroux, 2013] skewers our habits of ir/rational thinking, while claiming to remain optimistic about our rational prospects. I was equally unconvinced, or convinced of the opposite by the effectiveness of his own arguments.) While another impressive feature of this book is the comprehensiveness of its coverage of the most important relevant research, both empirical and philosophical, as well as the detailed examination of all of it, the author himself notes again and again that he is passing over some legitimate questions (in the interests of producing a finite book); and even acknowledges that some of the key issues he does raise may be insoluble.

As an example of Pölzler's treatment of his subject matter, here is his own synopsis of how he will go about discussing one of the main arguments in the moral realism debate, namely, Jesse Prinz's argument that morality is not (objectively) real because moral judgments are constituted by emotions:

In analyzing and assessing emotionism [the empirical premise of the argument], I will proceed in the by now familiar way. First, I will clarify the hypothesis's main concepts and explain in which sense it would have to be true in order for Prinz's argument to work. Then I will go on to evaluate the hypothesis by reference to existing [scientific] studies (in particular, studies that Prinz himself appealed to). It will turn out that, contrary to Prinz, these studies fail to support emotionism. Most of them lack in external or internal validity. Moreover, they would not even support a sufficiently close association of moral judgments and emotions if they were fully valid. I will close with some recommendations as to how emotionism could be tested in more valid ways (132).

There are scores of meticulous critiques to be found in this book, but the gist is simple enough. Take the main question: Is morality real? Obviously for any scientific research bearing on this question, yay or nay, to be effective, the study must be rigorous in its analysis of what morality and being real are or mean. A study of cat behaviour is not going to tell you about dogs, or at least you could not be sure that it does if you never studied dogs. Pölzler's conceptual critique is therefore that scientists are hardly to be relied on, as a rule, to get this right. Furthermore, even when they do, there is the further question of whether they operationalize the notions correctly; and here too Pölzler finds them often lacking. Note, then, that Pölzler critiques both the philosophy and the science in scientific work, as well as philosophers' often uncritical reliance on it (227).

Pölzler's most general suggestion for this collaboration is to strive for so-called reflective equilibrium between the empirical/scientific and conceptual/philosophic

investigations. A truly adequate analysis of, say, morality is not to be hoped for without scientific research into what people actually have in mind when they use that terminology; but, as just noted, that research cannot hope to elucidate the matter if it fails even to touch upon some of the possible meanings of morality. To progress on the question, therefore, both science and philosophy need to maintain a dialogue in order to stay mutually informed, which will result in continuing modifications to their respective analyses and protocols.

This proposal is eminently reasonable. But there remains a deeper problem. It may simply be the nature of the beast that notions like morality and being real will forever resist a univocal meaning (both among philosophers and lay folk). This is why I find it is almost tragicomic that the book ends with the author's detailed rehearsal of why his investigation has sown seeds of hope. For me this is a case of "the more he protests ...". (On page 184 Pözlér uses the phrase 'to advance the debate' to express the purpose of his positive suggestions. Presumably he means 'toward an ultimate solution'. But I cannot help but read the phrase cynically to mean, literally, to advance *debate*. Sometimes while reading this book the image formed in my mind of Pözlér coaching *both* pugilists in a boxing match: what one wants is a good fight, regardless of who wins or it's a draw.) But I do not doubt Pözlér's sincerity for a moment, nor do I fault the book in any way. Pözlér has made the best case that could be imagined for his optimism.

I would only suggest that one might recognize that there is an alternative to this optimism other than despair. For me the 'consolation of philosophy' has more and more become the opening up of a vast realm of humanistic analysis. Or really, I see two main roles for the philosopher: one critical, but like the Socrates of the early Dialogues, for whom there is and can be no resolution other than the wisdom of knowing that one knows nothing about the matters of most concern; the other humanistic, in the manner of literary and psychoanalytic investigations of the human (and animal) condition and self. What can make these still philosophical, in my preferred use of the term, is the effort to maintain a rational outlook on everything, however doomed to ultimate failure this project may be. I also believe both types of activity can offer practical benefits, such as ridding us of misplaced confidence due to confused thinking that can cause mischief in the world or our own life, and enhancing our imaginative capacity to care about others. But a final Practical Guide to Life (or a final answer to the question of moral objectivity) will prove eternally elusive.

Joel Marks  
University of New Haven  
Yale University

Maura REILLY. *Curatorial Activism: Towards an Ethics of Curating*. New York: Thames and Hudson, 2018. 240 pp.

The question of ethics is increasingly becoming an important aspect of curatorial practice and discourse. Curatorial ethics manifest in various ways: there are ethics of

museum spaces, ethics of exhibition-making, ethics of acquisition and collecting, the rights of artists, discourses around censorship, and the ethics of curators, curatorial responsibilities and so on. Central to most of these issues is the ethical dilemma between subject and object. Does the ethics of curating lean towards the objects of curation or the subjects of curation? In a simpler vocabulary, is the curator's ethics guided towards the artists or the art works?

Maura Reilly's *Curatorial Activism: Towards an Ethics of Curating*, is a welcome opening into the ethical facets of curating focused on the subject, based primarily on identity politics of the artist. A recent review of the book by Emma Mahony begins with an element of surprise at the realization that the world is still a heavily lopsided and unequal place for the arts. Given that Mahony is party to the discourse of the global North, such surprise is not unusual. However, as Mahony seems to claim, Maura Reilly's book does not *break any myths* because in the global 'South' which includes the marginalized, the coloured, the feminist, the queer, the Dalit and others, such questions are still in formation, and very much an everyday encounter. As a reader and curator from the 'Other' world (to borrow the author's terminology), reading Reilly's book at its best, offers an explication of the obvious conundrums and contradictions glaring in the face of the practical world of contemporary arts. However, its significance lies in the fact that it provokes questions of ethics to be considered and extended by scholars and critics of the 'Other' world with utmost urgency.

The author, who has been curator at the Elizabeth A. Sackler Center for Feminist Art at the Brooklyn Museum, focusses the book on three issues – gender, race and sexuality, all of which have identity politics at their core. These issues rise agonistically against the inherently mono-dimensional Euro/US-centric art historiography, infamously proposed by Alfred H. Barr with clearly defined 'isms' that reinforce the figure of heterosexual, white, male genius. Reilly is quick in dismantling this history by churning out numbers that expose appalling levels of discrimination (18). Citing the *1986 Report Card* by The Guerrilla Girls alongside the *2015 Report Card* by Pussy Galore, the author reiterates the inability of the art world to include 'Other' voices – "[...] the art world has not yet fully incorporated diverse or Other voices into the larger discourse – except, of course, as 'special' (read separatist) exhibitions such as Latin American Art, Women Artists, Islamic Art, African Art, and so on" (21). Influenced by scholars on racism and postcolonialism, Reilly's own practice has engaged with re-examining art historical canon's Euro/US-centrism and towards ensuring the rightful representation of the silenced, the 'doubly colonised', the unseen and unheard. In this book, Reilly offers insights into possible counter-hegemonic strategies by studying group exhibitions, specifically those that she finds "Talmudic, Wikipedia like" (15) in approach, enabling a wider and expansive discussion.

Reilly observes three 'strategies of resistance' that have been used by curators in exhibitions that attempt to overthrow these problems of unequal representations: Revisionism, Area Studies and Relational Studies, each with their unique intentions in destabilization of the canon. Revisionism strives to relook at certain histories to find gaps or failures that would ideally give a more complete sense of the concerned history and thereby illuminate it or 'improve' it. However, revisionism risks the presupposition

of the binary of the centre and the margin. Revisionism's focus on recuperative stances takes away possibilities of radical eschewing or rejection of histories. Its accommodative nature, thus, is often its undoing. This makes the choice of the histories to be revised very significant as a political strategy within revisionism.

The second strategy of the Area Studies focusses on work based either topographically or psychologically or culturally in particular orientations or specific regions. This includes race, geography, gender, sexual orientation etc. While area studies ensure rightful representation of marginalized groups, they risk becoming ghettoized into fixed identities, running the risk of the 'single story' that Chimamanda Ngozi Adichie speaks about (103). However, she points out the relevance of such specialized collections in the global context of prevalent inequality – "We cannot claim to live in a post-queer world when in some countries being queer, gay, bisexual, or transgender is punishable by death and in many more it is a criminal offence" (27). The third strategy is that of Relational Studies, which places emphasis on polyphonic registers and a flattening of definitive hierarchies to address a transnational condition of the world. In Reilly's words, "[...] it is a fundamental redefining of art practice, transnationally" (30). Heterogeneity and incoherence mark this kind of curating, which, in Reilly's words are 'writerly' exhibitions. The relational approach is thus considered by Reilly to be a fitting approach to an expanded 'transnational vision' of artistic production.

Throughout the book, Reilly chooses to address the binary of the white, male, heterosexual world and the rest of the world by calling the artists of her concern the 'Other artists'. Reilly's choice to address the artists of her concern as 'Other' artists, though rhetorical in intention, appears to the 'Other' reader with a hint of condescension. The Other has been fundamental to ethical pondering as seen in philosophers like Levinas. We see how the Other is constituted by its lack of intelligibility, thereby making it inevitable that for the preserving of the otherness of the Other, a failure of understanding is essential. However, one observes how the Same and Other also fall prey to binary structural politics. Post structuralist and South Asian philosophy, on the other hand, work with the idea of the multiplicity. Identity, in such frameworks, need not always adhere to an insider-outsider criterion, but is a fluid notion that, by understanding the contingent nature of the world, offers ways out of the Otherness rhetoric. This calls into question the prerogative of the white self-conscious west to consider the coloured non-west as the Other. Why is the coloured world not interested in calling the white world its Other? We realise that in attempting to answer these questions, one's underlying biases, presumptions and privileges are revealed. One might thereby understand that to even be able to speak of the Other as a *normalised* experience or encounter means to have internalized the privilege of selfhood, identity and a capacity of generosity to be dispensed with. In this regard, the form of the book, unfortunately, replays the strategies of Othering the 'Other' in not emancipating the signifying word from its history. With this small critical note in mind, reading Reilly's book is revealing in its expanse, scope, and possibilities that it throws out to the world to take up.

The author's first concern is with key praxiological questions pertaining to issues of affirmative action, visibility and minority quotas when it comes to gender. Reilly

outlines a survey of various exhibitions that could be considered as ‘catching up’ for women’s visibility across the UK, Europe, Canada and the US. These include exhibitions like *Women Artists: 1550 – 1950* curated by Linda Nochlin and Ann Sutherland Harris; *Bad Girls* curated by Kate Bush, Emma Dexter and Nicola White for Glasgow, Marcia Tucker for New York and Marcia Tanner for Los Angeles; *Inside the Visible: An Elliptical Traverse of 20<sup>th</sup> Century Art In, Of, and From the Feminine* curated by Catherine de Zegher.

The question of race becomes the author’s second concern. Several significant exhibitions since the late 1980s addressed the Eurocentrism within cultural institutions, whereby strategies such as inclusivity and fair representation were high on the agenda. Reilly notes that such notions of inclusion often turn into privileges of the male whites who become gatekeepers for the Other into their discourses as well as systems of white validation for token ‘native’ artists. In this context, citing Mosquera’s definition of curators as ‘mediators of cultural exchange’ (105), the author suggests a form of curatorial *modesty* as an asset to intercultural curatorial practices where curators resist claiming complete knowledge of the Other and can instead be open to collaboration with the ‘regional specialists’ to curate ethically. The author cites the much discussed *Magiciens de la Terre* curated by Jean-Hubert Martin in 1989 as an example of an exhibition (precursor of the ‘global exhibitions’ in contemporary art) that was criticized for its failures in doing justice to intercultural ethics. The other exhibitions she explores include *The Decade Show: Frameworks of Identity in the 1980s* curated by Julia P. Herzberg, Sharon F. Patton, Gary Sangster, Laura Trippi; *Mining the Museum* curated by Fred Wilson; *The Whitney Biennial* curated by Thelma Golden, John G. Hanhardt, Lisa Phillops, Elisabeth Sussman, and Jeannette Vuocolo; *Venice Biennale 2015: All the World’s Futures* curated by Okwui Enwezor and others. One notices that there is a higher rate of failure than success in these kinds of exhibitions that can be attributed to the large gap between intentions and the outcomes of the shows.

The third concern of the author is the history of censorship or ‘curatorial malpractice’ that is prevalent in the history of LGBTQ-concerned arts practices. In many of these cases, the author illustrates, through instances, how censorship and sins of omissions are rife in these curatorial ethics quoting Stern who argues that “[...] museums have a responsibility to acknowledge and consider the sexuality of artists in their collections when it is relevant to the work they are displaying [...] ignoring orientation amounts to curatorial malpractice” (162). Noting how even within the LGBTQ framework there are gaping exclusions like that of transgender artists, Reilly argues that there is still a long way to go. The exhibitions she considers include *Great American Lesbian Art Show (Galas)* curated by Terry Wolverton, Tyaga, Jody Hoeningner, Bia Lowe, Louise Moore, Barbara Stopha; *Extended Sensibilities: Homosexual Presence in Contemporary Art* curated by Dan Cameron; *Witnesses: Against Our Vanishing* curated by Nan Goldin; *In a Different Light: Visual Culture, Sexual identity, Queer Practice* curated by Nayland Blake and Lawrence Rinder; *Art Aids America* curated by Jonathan Katz and Rock Hushka and others.

Reilly treads a sensitive line between activism and ethics in her book, by placing it within her experience as a curator and theorizing towards ethical precepts from these

experiences. In conclusion, Reilly calls for a corrective strategy for change, taking on an activist's voice and arguing for the need of corrective, subversive and smart curatorial moves rather than curating resorting to revisionary affirmative politics. Questions of privilege, geo-political hierarchy, legacy, personal histories and responsibility become central to Reilly's proposition for curatorial ethics, one which several arts practitioners today are striving to address in their work. Key questions such as that of gallery representation and art collectors, media representation, the art market and economy, the constitution of boards, directors and curatorial committees and processes of selections, acquisition-making and valuation criteria become important to consider in this project. The curatorial activist is thus imagined as one whose intention is the addressing of various such questions and is sensitive to the ethical way of acting in this world. In her attempt at contributing to the 'creation of a just art world', Maura Reilly's work is a welcome precedent for 'Other' practitioners of contemporary curating to reflect incisively into our local, regional and particular ethical paradigms and bring out our stories to add to this discourse.

Srajana Kaikini  
Manipal Academy of Higher Education

Kieran SETIYA. *Midlife: A Philosophical Guide*. Princeton, NJ: Princeton University Press, 2017. 200 pp.

I have just turned 34. From the outside my life seems pretty okay. Yet, something does not feel quite right. My achievements mean less to me than I thought they would. When I think about the future, anxiety is a familiar companion. Gone the idealism. Gone the sense that I can be whatever I want to be. Am I having a midlife crisis (already, at this tender age)? And if yes, what can I do about it?

Looking for insights, I was delighted to find that my potential ailment has recently been taken up by philosophy. This is due, among others, to the pioneering work of MIT professor Kieran Setiya. In *Midlife: A Philosophical Guide* (marketed as a 'self-help book with a difference') Setiya analyses the midlife crisis and develops a number of suggestions about how to cope with it. The book is short and addressed to a general audience. Yet, in addition to finding helpful advice regarding their personal problems, academic philosophers' need for insightful analyses, arguments and distinctions is also well-served.

The book's first chapter provides a historical lesson. Having been introduced by psychoanalyst Elliott Jaques in 1965, the idea of the midlife crisis has become well-established in Western culture. Who hasn't heard jokes about or explanations in terms of such crises – about men (yes, typically men) who quit their well-paid jobs, start lusting after young women or buy expensive sports cars? Scientifically, Setiya's case for the midlife crisis as a distinct psychological condition is mainly based on recent studies in economics. These studies suggest that wellbeing is U-shaped. Being highest in young adulthood, it becomes low in middle-ages, and then again higher in old age.

What is the nature of this crisis in middle age? And how can it be alleviated? The distinctive feature of Setiya's book is its philosophical approach. Having been in the grips of a midlife crisis himself, he first introspectively accesses and presents what it was like for him. In this way he comes up with five main symptoms. By analyzing the concepts and inferences that give rise to these symptoms, Setiya shows that they mostly result from irrationality. He then suggests more appropriate ways of thinking that are supposed to function as a remedy against the midlife crisis.

The first purported symptom of midlife crises is illustrated by the example of John Stuart Mill. Early in his life Mill suffered from depression. He realized that even if his plans for profound social reforms succeeded, this would not bring him any significant joy or happiness. In Setiya's view, such ruminations may be cured by revising or expanding one's goals. Mill was always admiringly aware that one ought to care for things other than oneself. Those who directly strive for happiness will not achieve it. What he only discovered later was that solving problems, overcoming obstacles, and satisfying needs cannot be all there is to life either. In addition, one must also find a place for existential value: for things that make life positively good (such as sports or spending time with friends; or, in Mill's case, appreciating poetry).

A second way in which midlife crises are said to manifest themselves is in the feeling of missing out. When we are young, life abounds with options. We may well have chosen a different profession, a different spouse or a different place to live. At some point in time we realize – again, painfully – that many alternatives are no longer available. One way of reconciling ourselves with this limitation is to bear in mind that it is actually the flipside of something good. Life offers so many incommensurable values that we cannot possibly realize all of them. In addition, Setiya also warns against exaggerating the value of having options, and against romanticizing youth (which is also characterized by a lack of sense of identity and orientation).

Regretting past mistakes, misfortunes and failures is Setiya's third symptom of midlife crises. Parents may ease this symptom by bearing in mind that the bad things that happened prior to the conception of their children were a necessary condition for these children's existence. They have thus enabled great value. The book also raises awareness for the fact that alternative personal histories may have unfolded far less smoothly than we are inclined to think. In a different profession, with a different spouse or in a different place I might well be worse off than I am now. Finally, Setiya urges us to consider the details of what makes our actual lives good. This will render it rational for us to prefer them over most abstract 'what if?' scenarios (on which we ought not to ponder too much).

As we age we become increasingly aware of our mortality. The fourth manifestation of the midlife crisis is fear of death. Here philosophy is stretched to its limits. Setiya does not purport to show that this fear is wholly irrational. But there seem, nevertheless, to be ways of alleviating it. One of these ways involves bearing in mind that before our birth we were in a state of non-existence too. Although death deprives us of value, it is therefore no worse than what we have already been through (given that we consider past and future deprivations to be equally bad). Moreover, at least to the extent that it

is based on a desire to live forever, our fear of death actually turns out to be irrational. This is because, like the wish to be able to fly, such a desire cannot possibly be fulfilled.

Fifth and finally, Setiya characterizes the midlife crisis in terms of feelings of emptiness. These feelings are claimed to be due to a “[...] structural flaw in the pursuit of projects” (157) (following Arthur Schopenhauer). As long as we are working on a project we suffer from its incompleteness. But once we are done, things do not get better. We have lost what has given our life meaning and orientation. Setiya argues that the resulting feeling of emptiness can be mitigated by focusing more strongly on ‘atelic’ activities: activities that do not aim at some final state. For example, we could take a walk, spend time with friends or appreciate nature. By focusing on their process (e.g. doing philosophy) rather than their product (e.g. a completed book manuscript), even our engagement with projects can be given a healthy atelic touch. In Setiya’s view, mindfulness meditation is an effective training for achieving such a shift of focus.

This, then, is how the book characterizes and proposes to remedy the midlife crisis. Is it convincing? In answering this question we need to keep Setiya’s intentions in mind. *Midlife: A Philosophical Guide* uses philosophy in the service of cognitive psychotherapy. In other words, it does not only attempt to develop strong arguments; it also and primarily wants people to be better off (as a result of considering these arguments). For non-philosophers, the bar seems to be set quite high. Several parts of Setiya’s reasoning may be too abstract, sophisticated or long-winded for them. If you are able and willing to follow the book, however, I believe that it really can make a difference.

Given the book’s shortness and target audience, philosophers will have an easy time pointing out limitations and unsubstantiated assumptions. For example, why did Setiya not address regret for mistakes that harmed others (Christopher Hamilton. “Midlife: A Philosophical Guide, by Kieran Setiya.” *Mind* [2018; forthcoming]) or incorporate lessons from existential philosophy? Within these constraints, however, most of Setiya’s arguments strike me as plausible; and especially his suggestions about how to revise our thinking. My main worry mostly arises from the book’s consideration of scientific and introspective evidence. After reading it I am still unsure whether there is such a thing as the midlife crisis at all; and if there is, whether this crisis is most appropriately characterized by the five symptoms that Setiya focuses on.

First, is the midlife crisis a real psychological condition? Above I explained that Setiya’s main evidence for this claim derives from economic studies that suggest a U-curve in wellbeing. But this evidence is weak. In most studies wellbeing-drops in middle age are below 0.5 on a 7-point scale. This effect has been shown to be highly sensitive to a number of subtle methodological decisions. For example, plausible changes in control variables and models make it disappear completely (e.g. Paul Frijters and Tony Beaton. “The Mystery of the U-shaped Relationship Between Happiness and Age.” *Journal of Economic Behavior & Organization* 82/2-3: 525-542 [2012]). Some other studies have found that wellbeing remains stable across age (e.g. David G. Myers and Ed Diener. “Who is Happy?” *Psychological Science* 6/1: 10-19 [1995]) or that its U-shape is peculiar to developed countries in which the elderly are particularly well off (e.g. Angus

Deaton. “Income, Health, and Well-Being Around the World: Evidence from the Gallup World Poll.” *Journal of Economic Perspectives* 22/2: 53-72 [2008]). Finally, the results of any such research are contingent on controversial philosophical assumptions (such as that wellbeing can be measured subjectively), and have been found to be strongly influenced by the particular wording of questions (Arthur A. Stone, Joseph E. Schwartz, Joan E. Broderick, and Angus Deaton. “A Snapshot of the Age Distribution of Psychological Well-Being in the United States.” *PNAS* 107/22: 9985-9990 [2010]) (for a helpful overview see Beatriz Fabiola López Ulloa, Valerie Møller, and Alfonso Sousa-Poza. “How Does Subjective Well-Being Evolve with Age? A Literature Review.” *Journal of Population Ageing* 6/3: 227-246 [2013]).

Suppose Setiya is right that our wellbeing hits bottom in middle age. Is this finding best explained by what he takes to be the midlife crisis (i.e. by the sense that this cannot be all there is; the feeling of missing out; regret for mistakes, failures and misfortunes; fear of death; and a feeling of emptiness)? Answering this question is exacerbated by a lack of conceptual clarity. For example, one is left wondering how many of the above symptoms need to be present, and to what extent, in order for persons to be justifiably diagnosed with midlife crises. Setiya also does not clearly specify what he understands by ‘midlife’. In fact, his usage of the term seems too liberal. We are told that Camus suffered from a midlife crisis at the age of 30, and de Beauvoir at age 55 (which is still said to be ‘in the heart of middle-age’, suggesting that this age ranges even farther) (104).

The book involves several arguments for the above mentioned symptoms being most prevalent in middle age. But these arguments are very cursory and often depend on empirical assumptions that are not backed up by evidence. For instance, why should people be more afraid to die in middle age than in other periods of their lives? Here is Setiya’s answer: “In middle age, the limited span of human life is no longer an abstraction. You know from the inside what a decade means; those that remain to you can be counted on one hand. That can be a source of angst” (105). Sure, one is tempted to reply, but would we not expect people in old age to be even more afraid, then? After all, they have witnessed *several* decades passing by, and are much closer to death knocking on their doors.

Note also that if there really is a U-curve of wellbeing it might be plausibly explained in terms other than that of midlife crises as well. Setiya himself approvingly cites a study in which zoo workers rated the wellbeing of great apes (Alexander Weiss, James E. King, Miho Inoue-Murayama, Tetsuro Matsuzawa, and Andrew. J. Oswald. “Evidence for a ‘Midlife Crisis’ in Great Apes Consistent with the U-shape in Human Well-Being.” *PNAS* 109/49: 19949-19952 [2012]). Again, a slope in middle age emerged. This finding suggests that lifespan variation in wellbeing may (partly) have direct biological causes (rather than misguided thoughts about existential issues) (see also David Lykken, and Auke Tellegen. “Happiness is a Stochastic Phenomenon.” *Psychological Science* 7/3: 473-485 [1996]). Several other plausible alternative explanations of the U-curve – ranging from stress to unmet aspirations – would have to be ruled out as well.

My point here is not that Setiya is wrong in suggesting a midlife crisis that typically manifests itself in one or more of the above five symptoms. In the end his experiences

may well turn out to be representative (at least of some significant subset of the population, such as well off people in developed countries). I would rather like to draw attention to the fact that this claim can only be assessed by systematic empirical investigations. For example, experimental philosophers or psychologists could conduct studies that ask people at several points in life whether they have a sense that this cannot be all there is, a feeling of missing out, regret for mistakes, and so on; and experienced psychotherapists could report their perceptions about lifespan variations in these variables.

My own hunch is that Setiya's five symptoms are not (or at least not significantly) more prevalent at any particular period of adulthood. If this is right, then he should not have labelled his discussion as being about the 'midlife crisis'. Yet, it is important to note that this problem would take only little away from what makes his book so valuable. I take it to be highly plausible that almost all adults at almost all stages of their lives are prone to having misconceptions about the intricate existential issues that *Midlife: A Philosophical Guide* addresses. Setiya does an excellent job in revealing and correcting some of these misconceptions. His book may therefore be helpful to even more people than he thought. Moreover, it demonstrates how fruitfully contemporary analytic philosophy can inform cognitive psychotherapy (a link that deserves far more attention than it has so far received).

I started this review by lamenting my own personal crisis. Having read Setiya's book, I have become fairly certain that it is unconnected to my age. At its roots, however, I do find some of the very misconceptions that the book corrects (such as an underappreciation of the richness of life and an excessive focus on projects). By enabling me to tackle these misconceptions, to me the book has succeeded in making a real difference.

Thomas Pölzler  
University of Graz

---

## Contributors

---

FRIEDEMANN BIEBER is a doctoral student in Political Philosophy at the University of Zurich, funded by the Swiss National Science Foundation (SNSF). The monograph he is working on is provisionally entitled *Systemic Risk and Public Goods*. He holds an MA in Philosophy from the University of Zurich and previously read Philosophy, Politics and Economics at the University of Oxford.

Email: [friedemann.bieber@philos.uzh.ch](mailto:friedemann.bieber@philos.uzh.ch)

EIKE DÜVEL is a PhD candidate at the Karl-Franzens-University Graz and member of the FWF Doctoral Program Climate Change. His research interests are in the ethics of risk imposition, especially contractualist theories of risk imposition, and in foundational and methodological questions in practical philosophy more generally. He is currently working and publishing on the fair distribution of risk in society and on the theory of legitimate expectations in the context of climate change. His latest publication is “Human Rights-Based Precautionary Approaches and Risk Imposition in the Context of Climate Change” (2018).

Email: [eike.duevel@uni-graz.at](mailto:eike.duevel@uni-graz.at)

JEFF HAINES is a doctoral candidate in philosophy and political science at Florida State University. He has taught both philosophy and political science courses, including on applied ethics, critical thinking, game theory, and political psychology. He holds an MA in Humanities and an MS in Political Science. His research interests are in immigration, global justice, risk, moral psychology and game theory.

Email: [jhaines@fsu.edu](mailto:jhaines@fsu.edu)

SVEN OVE HANSSON is professor in philosophy at the Royal Institute of Technology, Stockholm. He is editor-in-chief of *Theoria* and of the two book series *Outstanding Contributions to Logic* and *Philosophy, Technology, and Society*. His research on risk and safety focuses on ethical and epistemological aspects. His other philosophical research includes contributions to the philosophy of science and technology, decision theory, logic, and moral and political philosophy. He is member of the Royal Swedish Academy of Engineering Sciences and past president of the Society for Philosophy and Technology. He

has published about 350 papers in refereed international journals and books. His books on risk-related topics include *The Ethics of Risk* (2013), *Social and Ethical Aspects of Radiation Risk Management* (with Deborah Oughton, 2013), *The Role of Technology in Science* (2015), *The Ethics of Technology. Methods and Approaches* (2017), and *Handbook of Safety Principles* (with Niklas Möller, Jan-Erik Holmberg, and Carl Rollenhagen, 2018).  
Email: soh@kth.se

MADELEINE HAYENHJELM is a Senior Lecturer in Philosophy at Umeå University, Sweden, with ethics of risk as one of her main research interests. She is currently engaged in a research project on the ethical implications of CRISPR/Cas9. When not researching she teaches Ethics and Philosophy of Risk. She received her PhD from KTH, Stockholm, for a thesis on risk communication and trust, and she held a three-year post-doc position at University College London on the ethics of risk, and a one-year position as Lecturer in Ethics at Edinburgh University, before joining the department at Umeå University. Relevant publications include: “Compensation as Moral Repair and as Moral Justification for Risks” in *Journal of Ethics, Politics, and Society* (forthcoming), and “The Moral Problem of Risk Impositions: A Survey of the Literature” (2011), co-written with Prof. Jonathan Wolff, in *European Journal of Philosophy*.  
Email: madeleine.hayenhjelm@umu.se

NIMROD Z. KOVNER is a post-doctoral research fellow in the ‘Human Rights Under Pressure’ Interdisciplinary Research Program of The Minerva Center for Human Rights at the The Hebrew University of Jerusalem. He wrote his doctoral dissertation at the Government Department of the London School of Economics and Political Science (LSE) and holds a BA in Philosophy, Politics and Economics and an MA in Political Science, both from The Hebrew University of Jerusalem. Dr. Kovner specializes in applied moral and political philosophy. He is interested in responsibility for aggregated harms and risks, and he works on issues such as climate change, immigration, public health ethics, and vaccination policy.  
Email: nimrod.kovner@mail.huji.ac.il

LUKAS H. MEYER received his doctorate from the University of Oxford, completed the habilitation at the University of Bremen and has held postdoc positions in the USA (at Harvard University and Columbia University). Before becoming a Professor at the University of Graz in 2009, he was an Assistant Professor at the University of Berne. Lukas Meyer’s research has been mainly in the area of intergenerational justice, understood to include the relations presently living people have to past as well as future people. His other areas of research are: international justice, normative authority and legitimacy, and non-ideal theorizing. His publications include *Historical Emissions and Climate Justice* (with Pranay Sanklecha: CUP 2017); *Intergenerational Justice* (with Axel Gosseries: OUP 2009);

## CONTRIBUTORS

(2008): “Intergenerational Justice” in *Stanford Encyclopedia of Philosophy* 2008); *Historische Gerechtigkeit* (de Gruyter 2005). Meyer was one of the first two philosophers to serve as lead author for the next Assessment Report of the IPCC (International Panel on Climate Change). He is a founding editor of *Moral Philosophy and Politics* (de Gruyter).  
Email: lukas.meyer@uni-graz.at

ADRIANA PLACANI is Assistant Professor in Political Philosophy at the University of Graz, Austria. Previously, she was Postdoctoral Research Fellow at the Centre for Advanced Studies in Sofia, Bulgaria, and Visiting Scholar at Columbia University in New York. She received her PhD from LUISS Guido Carli in Rome, Italy. Her work focuses on the ethics of risk, with particular emphasis on the morality of risk impositions and responsibility for risks. Relevant publications include “When the Risk of Harm Harms” (in *Law and Philosophy*) and “Ethical Dimensions of Risk: Beck and Beyond” (in *Law in the Risk Society*).  
Email: adriana.placani@uni-graz.at

THOMAS PÖLZLER is researcher and lecturer at the philosophy department of the University of Graz, Austria. His main areas of research are moral psychology and meta-ethics. He is also interested in the philosophy of Albert Camus, environmental ethics, intergenerational justice, and basic needs. Dr. Pölzler is the author of *Moral Reality and the Empirical Sciences* (Routledge, 2018). His articles have been published in journals such as *Synthese*, *Review of Philosophy and Psychology*, and *Journal of Value Inquiry*.  
Email: thomas.poelzler@uni-graz.at

KLAUS STEIGLEDER is Professor of Applied Ethics at the Faculty of Philosophy and Education, Ruhr University Bochum, Germany. He is director of the interdisciplinary master’s program Ethics – Economics, Law, and Politics and member of the board of the Research Department Closed Carbon Cycle Economy at Ruhr University. His work focuses on the justification of normative ethics and the theory of rights, ethics of climate change, risk ethics, energy ethics, and economic ethics.  
Email: klaus.steigleder@ruhr-uni-bochum.de

HARALD STELZER is Professor of Political Philosophy at the Department of Philosophy of the University of Graz. His main research interests are in the normative evaluation of emerging technologies with a special focus on climate engineering and internet and communication technologies as well as in political philosophy on the liberalism-communitarianism debate and the philosophy of Karl Popper. Recent publications include a book on the critique of the moral philosophy of communitarianism (2016) as well as articles on climate-engineering in *Environmental Values* (2016) and *Annals of the University of Bucharest - Philosophy Series* (2017).  
Email: harald.stelzer@uni-graz.at

AVIEZER TUCKER is a Research Associate at Harvard University. Previously, he was Assistant to the Director of the Energy Institute at the University of Texas in Austin, and held a Research Fellowship on democratic evaluation of new technologies at the Research School of Social Sciences of the Australian National University in Canberra. He holds a BA in History (Tel Aviv University), and PhD in Philosophy (University of Maryland College Park) and a postdoc in Political Science (Columbia University). His research concentrates on Political Philosophy and Theory, Epistemology and the Philosophy of Science, Science Studies, and their numerous intersections. His main publications include *The Legacies of Totalitarianism: A Theoretical Perspective* (Cambridge University Press 2015); *Our Knowledge of the Past A Philosophy of Historiography* (Cambridge University Press 2004); *Panarchy: Political Theories of Non-Territorial States* (Routledge 2015); and *A Companion to the Philosophy of History and Historiography* (Blackwell, 2008). He has published numerous articles and reviews on politics and energy in *The American Interest* and *Foreign Affairs*, as well as in academic journals of political philosophy, epistemology, and philosophy of science.

Email: [avitucker@yahoo.com](mailto:avitucker@yahoo.com)